

Fri Mar 12 10:35:45 2004

us-10-084-817-30.fst

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 10, 2004, 09:18:38 ; Search time 6129.06 Seconds

(Without alignments)  
18036.975 Million cell updates/sec

Title: US-10-084-817-30

Perfect score: 3702

Sequence: 1 gcctaggggaatcagcagcat.....aaggggcggttaccgagaac 3702

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :  
1: em\_estba:\*  
2: em\_estbma:\*  
3: em\_estma:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_estc:\*  
9: gb\_estl:\*  
10: gb\_est2:\*  
11: gb\_estc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estcom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vit:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rtd:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vrl:\*  
28: gb\_gss1:\*  
29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	822.4	22.2	923	13	BQ718603 AGENCOURT
2	809.8	21.9	1201	9	AL577531 AL577531
3	790.6	21.4	826	14	CK229986 ILUMINIGEN
4	769	20.8	941	13	BQ925009 AGENCOURT

5	768.2	20.8	846	13	BUI51076 AGENCOURT
6	768	20.7	881	14	CD359921 AGENCOURT
7	753.8	20.4	827	14	CD643635 AGENCOURT
8	747.4	20.2	749	9	AL700988 DKEP2686D
9	737.8	19.9	823	14	AUI32922 AUI32922
10	732	19.8	823	14	CD656598 AGENCOURT
11	725.4	19.6	775	12	B0010124 AGENCOURT
12	719.6	19.4	775	9	A1798679
13	708.8	19.1	753	9	A1926727
14	708.8	19.1	753	9	A1926727
15	701.2	18.9	735	13	BX642011
16	698.4	18.9	735	13	BX642011
17	692.6	18.7	699	12	CB998236 AGENCOURT
18	691.8	18.7	699	12	B0014967
19	688.8	18.6	704	13	BX472022
20	688.4	18.6	704	13	BG209023
21	681.4	18.4	834	13	BUS67576
22	680.8	18.4	887	10	BE749163
23	678	18.3	760	14	CB229386
24	671.2	18.1	727	13	BX504392
25	668.2	18.0	704	9	AUI39912
26	666.2	18.0	727	13	AUI39912
27	665.2	18.0	727	13	AUI39912
28	663	17.9	1201	9	AL554932
29	662.2	17.9	840	14	CB310135
30	662	17.9	730	13	BX504360
31	659.4	17.8	677	13	BX611260
32	654.6	17.7	675	13	BX481847
33	654.6	17.6	660	13	BX471556
34	650.4	17.6	662	12	B0014498
35	650.4	17.6	662	14	CA440470
36	650.2	17.6	655	13	BX476085
37	649.4	17.5	651	13	BX479376
38	647.8	17.5	664	13	BX505684
39	647.2	17.5	662	14	CA432157
40	645.6	17.4	662	13	B0625233
41	641.8	17.3	675	9	AUI39864
42	641	17.3	663	12	BM723358
43	639.4	17.2	851	14	CD251635
44	637.6	17.2	759	14	CB962709
45	636	17.2	646	14	CA771730

ALIGNMENTS

RESULT 1  
BQ718603  
LOCUS  
DEFINITION  
AGENCOURT\_8295719 Lupeki\_sympathetic\_trunk Homo sapiens cDNA clone  
IMAGE:6188951.5, mRNA sequence.  
BQ718603  
AGENCOURT\_8295719 GI:21857500  
EST.  
Homo sapiens (human)  
Homo sapiens  
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
NIH-MGC http://mgc.nci.nih.gov/  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, P.D.  
Email: cgabs-remail.nih.gov  
Tissue Procurement: Dr. James R. Lupeki  
cDNA Library Preparation: Life Technologies, Inc.  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/ILNL at:  
http://image.llnl.gov  
Plate: ILNL3585 row: 6 column: 24  
High quality sequence stop: 583.

## FEATURES

Source

Location/Qualifiers

```

1. 923
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6188951"
/sex="male"
/tissue_type="sympathetic trunk"
/dev_stage="adult, 16 yr"
/lab_host="DH10B"
/clone_lib="Lupski_sympathetic_trunk"
/ncbi_vector="PCMV-SPORT6 (Life Technologies); Site 1:
NotI; Site 2: SalI; cDNA made by oligo-dT priming.
Directionally cloned using the following adaptors:
5'-TCGACCCGAGGCTCG-3' and
5'-GACTAGTCTTCAATATGCGACGCGCGCCCT(15)-3'. Size selected >
1 kb for average insert length 1.9 kb. This is a primary
library, non-amplified. Library constructed by Life
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
College of Medicine); available through Life
Technologies."

```

## ORIGIN

Query Match 22.2%; Score 822.4; DB 13; Length 923;

Best Local Similarity 95.9%; Pred. No. 1.9e-132;

Matches 885; Conservative 0; Mismatches 34; Indels 4; Gaps 4;

```

QY 492 AGGAGATGCGACTCGAAGCAGAAAGATGCGCTACCGAAGAAATGACGATGATGAAGAGG 551
DB 1 AGGAGATGCGACTCGAAGCAGAAAGATGCGCTACCGAAGAAATGACGATGATGAAGAGG 60
QY 552 AGGAGATGCGACTCGAAGCAGAAAGATGCGCTACCGAAGAAATGACGATGATGAAGAGG 611
DB 61 AGGAGATGCGACTCGAAGCAGAAAGATGCGCTACCGAAGAAATGACGATGATGAAGAGG 120
QY 612 AAGATCCTTGGGACAGATGACCGACACAGATGAGTGAATGCCAAGACAGTGTGCTG 671
DB 121 AAGATCCTTGGGACAGATGACCGACACAGATGAGTGAATGCCAAGACAGTGTGCTG 180
QY 672 ACGAGAGAGCCCAAGACACACACACACACACACACACACACACACACACACACACAC 731
DB 181 ACGAGAGAGCCCAAGACACACACACACACACACACACACACACACACACACACACAC 240
QY 732 TCCTGAGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 791
DB 241 TCCTGAGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
QY 792 AGCGGCAAGAGAGTTCGACCCCAACATTAACATGCAATGCAATGCAATGCAATGCAAT 851
DB 301 AGCGGCAAGAGAGTTCGACCCCAACATTAACATGCAATGCAATGCAATGCAATGCAAT 360
QY 852 GAATGCAAAATGACACAGCAGAAATGAATGAATGAATGAATGAATGAATGAATGAATGTC 911
DB 361 GAATGCAAAATGACACAGCAGAAATGAATGAATGAATGAATGAATGAATGAATGAATGTC 419
QY 912 GCCAAGAAAGATGACGATGAGAGAAACGAAACGTCACCAAGTCTCTACCAAGAAATG 971
DB 420 GCCAAGAAAGATGACGATGAGAGAAACGAAACGTCACCAAGTCTCTACCAAGAAATG 479
QY 972 ATTGAGAGGATGCTGAAGAAACAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1031
DB 480 ATTGAGAGGATGCTGAAGAAACAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 539
QY 1032 AGAAGCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1091
DB 540 AGAAGCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 598
QY 1092 CAATGAAAGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1151
DB 599 CAATGAAAGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 658
QY 1152 AAGAGCTTAAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1211

```

```

DB 659 AAGAGCTTAAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 718
QY 1212 AGATGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1270
DB 719 AGATGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 778
QY 1271 GAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1330
DB 779 GAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 838
QY 1331 GAACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1389
DB 839 GAACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 898
QY 1390 AAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1412
DB 899 AAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 921

```

## RESULT 2

AL577531/c 1201 bp mRNA linear EST 01-JUN-2003

LOCUS AL577531 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA

DEFINITION clone CSOD1087YP20 3-PRIME, mRNA sequence.

ACCESSION AL577531

VERSION AL577531.2 GI:31315791

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 1201)

Li W.B., Gruber C., Jessee J., and Polayes D.

Full-length cDNA libraries and normalization

Unpublished (2001)

On Feb 16, 2001 this sequence version replaced GI:12940753.

COMMENT

Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr

Library was constructed by Life Technologies, a division of

Invitrogen. This sequence belongs to sequence cluster 4268.f For

more information about this cluster, see

http://www.genoscope.cns.fr/

cgi-bin/cluster.cgi?seq=CSOD1087DH10NPL&amp;cluster=4268.f. Contact :

Feng liang Email: fliang@life.techn.com URL :

http://fulllength.invitrogen.com/Invitrogen Corporation 1600

Faraday Avenue Genoscope sequence ID : CSOD1087DH10NPL.

Location/Qualifiers

1..1201

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="CSOD1087YP20"

/tissue\_type="PLACENTA COT 25-NORMALIZED"

/clone\_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"

/note="Test strand cDNA was primed with a NotI-oligo(dT)

primer. Five prime end enriched, double-strand cDNA was

digested with Not I and cloned into the Not I and EcoR V

sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 21.9%; Score 809.8; DB 9; Length 1201;

Best Local Similarity 83.1%; Pred. No. 2.7e-130;

Matches 942; Conservative 55; Mismatches 124; Indels 12; Gaps 8;

```

QY 2166 AAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2225
DB 1144 AAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1085
QY 2226 AACTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2285
DB 1084 AARGRGRGSGGGGAGCCGGSAAAYCTSGGGGGGGAGAGAGAGGGGKCMWAGAAAPAWTTA 1025

```

```

QY 2286 AAGCTGCTGAGAAAGCCGAGAGATGCGAGAGATGCGTGTGACATGACAGAAACCAT 2345
DB 1024 AAGGCGARBSAAGSYKCTAAACCCAGABATSCCAGAGTGGTATTMTWMAAAA 965
QY 2346 TCAAGT--GTTTACTCTTAAGGTTTCATCTTCAAGATGAAAGAGCGACGACATTTT 2402
DB 964 ACCATTCAAKTCTTAACCCCAARGTTC--TTTCCCAAMARARGSGRGCGA--TTT 909
QY 2403 TGAATAGTCTGTGCAAAAGAGAGTGTGTCATGCAATGCAACCATCAAGCAGATAGTCT 2462
DB 908 TTRAAAGTTTGGMGAAAGAGAG--KGTTCATTCMCCATCAACAGATAGTCT 850
QY 2463 CCAAGATTGACAGCAGACTGAGAGCATATACAGTGCATTTGAGGAAACAAAAGCCAA 2522
DB 849 CCAAGATTGACAGCAGACTGAGAGCATATACAGTGCATTTGAGGAAACAAAAGCCAA 790
QY 2523 AACCTCAAAAGCGGAGCGCTCGGATCTTCTGTTCTGTAAGGTGACGACATCA 2582
DB 789 AACCTCAAAAGCGGAGCGCTCGGATCTTCTGTTCTGTAAGGTGACGACATCA 730
QY 2583 AAGATATGTGGAGAAAGGAGATGTGTTTCATCCCACTGACGACGACACCAATA 2642
DB 729 AAGATATGTGGAGAAAGGAGATGTG--TTTCATCCCACTGACGACGACACCAATA 671
QY 2643 AAGAACTGTGCTGTTGAAGTATGAGGTTTCTAGCCGATCAATGAAATGGCTAATAA 2702
DB 670 AAGAACTGTGCTGTTGAAGTATGAGGTTTCTAGCCGATCAATGAAATGGCTAATAA 611
QY 2703 CCCCATGTGAAAACAATGCACTGCTCCCAACCTTCTGACTTGAGACGAGACGAT 2762
DB 610 CCCCATGTGAAAACAATGCACTGCTCCCAACCTTCTGACTTGAGACGAGACGAT 551
QY 2763 CCAGCAGGAGGAACTCTGAGGAAAGCAATCTGAGATGATGATCTCCCACTAAG 2822
DB 550 CCAGCAGGAGGAACTCTGAGGAAAGCAATCTGAGATGATGATCTCCCACTAAG 491
QY 2823 TTGAGACAGTTTCAGAAAGAACCCAAAGCTCAGAGACGAGACGAGTGTGAGAG 2882
DB 490 TTGAGACAGTTTCAGAAAGAACCCAAAGCTCAGAGACGAGACGAGTGTGAGAG 431
QY 2883 GCTAATTCGCTGCTTTTGTATTTATGTTATGTTATGTTATGTTATGTTATGTTAT 2942
DB 430 GCTAATTCGCTGCTTTTGTATTTATGTTATGTTATGTTATGTTATGTTATGTTAT 371
QY 2943 TTTTCATATCCAGTAAACCCATGATATATATCACTATATTAATCAAGTCTAGA 3002
DB 370 TTTTCATATCCAGTAAACCCATGATATATATCACTATATTAATCAAGTCTAGA 311
QY 3003 GATGTTCAATGTAAGTAACTGCTTTGACAGAGCCTGTTTAAAGAAACCCATGCT 3062
DB 310 GATGTTCAATGTAAGTAAACCTGCTTTGACAGAGCCTGTTTAAAGAAACCCATGCT 251
QY 3063 GTGAAATAGAGACTTTTCTACTGATCATATCACTGATCTGATGAGAGTATACCAAC 3122
DB 250 GTGAAATAGAGACTTTTCTACTGATCATATCACTGATCTGATGAGAGTATACCAAC 191
QY 3123 ACATCTGAATGACAGAAAGATCCAAGTTTAAATGCTGCGGAGATGTGCAATATCT 3182
DB 190 ACATCTGAATGACAGAAAGATCCAAGTTTAAATG--CTGCGAAATATMAMAMATAT 132
QY 3183 AGAAATAGAACCTAGTTTGTGTTTAAATACAGAAAT--CATGTTGTTCTGACT 3241
DB 131 AGAAATAGAACCTAGTTTGTGTTTAAATACAGAAATGAGATGTTGTTCTGACAG 72
QY 3242 TTTTAAATAGAGTGA--AGAAATATTTAGTAGGCAATGTACACTTTT 3293
DB 71 TTTTAAATAGAGTGA--AGAAATATTTAGTAGGCAATTTMMAACTTGT 19

```

RESULT 3  
CK229986 826 bp mRNA linear EST 09-DEC-2003  
LOCUS CK229986 826 bp mRNA linear EST 09-DEC-2003

DEFINITION ILLUMIGEN MCO 254 Katze\_MMP2 Macaca mulatta cDNA 5' similar to human CALD1 (Hs\_443811), mRNA sequence.  
ACCESSION CK229986  
VERSION CK229986.1 GI:39636090  
KEYWORDS EST.  
SOURCE Macaca mulatta (rhesus monkey)  
ORGANISM Macaca mulatta  
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea; Cercopithecidae; 1 (bases 1 to 826)  
REFERENCE Katze, M.G., Thomas, M., Korth, M., Iadonato, S.P. and Magness, C.L. Large-scale Rhesus Macaque cDNA Sequencing Unpublished (2003)  
AUTHORS CONTACT: C. Magness  
JOURNAL Illumigen Biosciences Inc.  
COMMENT 2203 Airport Way S, Suite 450, Seattle, WA 98134, USA  
Tel: 2063780400  
Fax: 2063780408  
Email: cmagness@illumigen.com  
Sequenced on 2003.09.19. 686 Q20 bases. Assemblies in contig w/ 1 member(s). Contig contains 1 (0%) 11b members.  
PCR Primers  
FORWARD: CCCTCACTAAAGGAGACAAA  
BACKWARD: CACTATAGGCGGATGCGTA  
Insert Length: 826 Std Error: 0.00  
Plate: C1000001 row: H column: 03  
Seq primer: CCCTCACTAAAGGAGACAAA  
POLYA=No.

## FEATURES

source  
1. 826  
/location/Qualifiers  
/organism="Macaca mulatta"  
/mol\_type="mRNA"  
/strain="Indian"  
/db\_xref="taxon:9544"  
/sex="male"  
/dev\_stage="newborn infant"  
/lab\_host="E. coli SOLR"  
/clone\_id="Katze\_MMP2"  
/note="Organ: Placenta; Vector: Uni-ZAP XR, site 1: Ecor I, site 2: Xho I; Created from StrataGene ZAP-CDNA Synthesis kit (catalog #200400) and ZAP-CDNA Gigapack III Gold Cloning Kit (catalog #200450)"

## ORIGIN

Query Match 21.4%; Score 790.6; DB 14; Length 826;  
Best Local Similarity 98.2%; Pred. No. 6.3e-127;  
Matches 810; Conservative 0; Mismatches 14; Indels 1; Gaps 1;  
QY 1934 GAATTCATGACCCCAAACTTAACATATCTGAGATCTTTCAAGCGCCCTGAGAGGAG 1993  
DB 1 GAATTCATGACCCCAAACTTAACATATCTGAGATCTTTCAAGCGCCCTGAGAGGAG 60  
QY 1994 GCCAGCGTGAACACCAAGAGAGCTGAGGCGCCCCCAGGTGAAGCCGCAAAAGCTG 2053  
DB 61 GCCAGCGTGAACACCAAGAGAGCTGAGGCGCCCCCAGGTGAAGCCGCAAAAGCTG 120  
QY 2054 GAGGAGCTTGCTGCTGCTGCGCGGGGAGACCGAGAGGAAGTTGAGAAAGCTCAACAG 2113  
DB 121 GAGGAGCTTGCTGCTGCTGCGCGGGGAGACCGAGAGGAAGTTGAGAAAGCTCAACAG 180  
QY 2114 AAGCAGAGAGAGCGGCTTTGAGAGCTGAGAGAACTCAAGAAAAGAGGAGAGAG 2173  
DB 181 AAGCAGAGAGAGCGGCTTTGAGAGCTGAGAGAACTCAAGAAAAGAGGAGAGAG 240  
QY 2174 AAGGCTCGGAGAGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2233  
DB 241 AAGGCTCGGAGAGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300  
QY 2234 GAGGAG 2293  
DB 301 GAGGAG 360

QY 2294 GAGAAACCCAGAGATGCGCAGAGATGGCTTGTGATGACAGAGAAACCATTCAGTGT 2353  
 Db 361 GAGAAACCCAGAGATGCGCAGAGATGGCTTGTGATGACAGAGAAACCATTCAGTGT 420  
 QY 2354 TTCCTCTTAAAGGTTCTATCTCTCAAGATGAGAGAGAGAGAGATTTTGAATAGTCT 2413  
 Db 421 TTCCTCTTAAAGGTTCTATCTCTCAAGATGAGAGAGAGAGATTTTGAATAGTCT 480  
 QY 2414 GTGCAGAAAGAGAGATGCTGTCAATGACCCATCAAGCAGATAGTCTCCAGATGAC 2473  
 Db 481 GTGCAGAAAGAGAGATGCTGTCAATGACCCATCAAGCAGATAGTCTCCAGATGAC 540  
 QY 2474 AGCAGACTGAGAGATGATACCACTGATGAGAGAGAGAGAGAGAGAGAGAGAG 2533  
 Db 541 AGCAGACTGAGAGATGATACCACTGATGAGAGAGAGAGAGAGAGAGAGAGAGAG 600  
 QY 2534 CCGGAGAGCTCGAGATCTTCTGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2593  
 Db 601 CCGGAGAGCTCGAGATCTTCTGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660  
 QY 2594 GAGAAAGGAGATGTTTCTATCTCCCACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2653  
 Db 661 GAGAAAGGAGATGTTTCTATCTCCCACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720  
 QY 2654 GGCTGAGAGTGGGGGTTTCTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2712  
 Db 721 GGCTGAGAGTGGGGGTTTCTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780  
 QY 2713 AAACAGTCACTGCTCCCAACCTTGTGATGAGAGAGAGAGAGAGAGAGAGAGAGAG 2757  
 Db 781 AAACAGTCACTGCTCCCAACCTTGTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 825

RESULT 4  
 LOCUS BQ925009 941 bp mRNA linear EST 20-AUG-2002  
 DEFINITION AGNCOURT 8840364 Lupski\_sciatic\_nerve Homo sapiens cDNA clone IMAGE:6205114 5', mRNA sequence.  
 ACCESSION BQ925009  
 VERSION BQ925009.1 GI:22340040  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 941)  
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-remail.nih.gov  
 Tissue Procurement: Dr. James R. Lupski  
 cDNA Library Preparation: Life Technologies, Inc.  
 DNA Sequencing by: Agencourt Bioscience Corporation (LML)  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LML at:  
 http://image.llnl.gov  
 Plate: LLM13627 row: h column: 11  
 High quality sequence stop: 617.  
 Location/Qualifiers  
 .941  
 location/Qualifiers

## FEATURES

Source  
 1..941  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:6205114"  
 /sex="male"  
 /tissue\_type="sciatic nerve"  
 /dev\_stage="adult, 70 yr"  
 /lab\_host="DH10B"  
 /clone\_lib="Lupski\_sciatic\_nerve"  
 /note="Vector: PCMV-SPORT6 (Life Technologies); Site\_1:  
 NotI; Site\_2: SalI; cDNA made by oligo-dT priming.

Directionally cloned using the following adaptors:  
 5'-TCAGCCAGCGCG-3' and  
 5'-GACTGTTCTAGTCCGAGCGCGCC(15)-3'. Size selected >  
 1 kb for average insert length 1.87 kb. This is a primary  
 library, non-amplified. Library constructed by Life  
 Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor  
 College of Medicine) and is available through Life  
 Technologies."

Query Match 20.8%; Score 769; DB 13; Length 941;  
 Best Local Similarity 93.9%; Pred. No. 3.4e-123;  
 Matches 821; Conservative 0; Mismatches 51; Indels 2; Gaps 2;  
 ORIGIN  
 245 CTGTCCTCTGTTAAAGAAATGATCTCTTCTGCACTTCTCTGGAAGAGATT 304  
 1 CTGTCCTCTGTTAAAGAAATGATCTCTTCTGCACTTCTCTGGAAGAGATT 60  
 Db 1 TCAGACTCAAGGTATCATTTGAGAAATTCAGATGATCAATGAAATTCACAGGAGATT 364  
 QY 305 TCAGACTCAAGGTATCATTTGAGAAATTCAGATGATCAATGAAATTCACAGGAGATT 120  
 Db 61 TCAGACTCAAGGTATCATTTGAGAAATTCAGATGATCAATGAAATTCACAGGAGATT 120  
 QY 365 GGTGACCAACAGAGAGCTCAGACATCTGATTTGCTGACCTGTCCAGACATCTGCT 424  
 Db 121 GGTGACCAACAGAGAGCTCAGACATCTGATTTGCTGACCTGTCCAGACATCTGCT 180  
 QY 425 CCCTGAACTGTAATTCACACCATGATGATTTGAGCGGTGAGAGAGAGAGAGAGAG 484  
 Db 181 CCCTGAACTGTAATTCACACCATGATGATTTGAGCGGTGAGAGAGAGAGAGAGAG 240  
 QY 485 AAGAGAGAGAGATGAGTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 544  
 Db 241 AAGAGAGAGAGATGAGTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300  
 QY 545 GAAAG 604  
 Db 301 GAAAG 360  
 QY 605 CAGAGAGAGAGATCTTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 664  
 Db 361 CAGAGAGAGAGATCTTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420  
 QY 665 GTGCTGAG 724  
 Db 421 GTGCTGAG 480  
 QY 725 GCGCATTTCTGAG 784  
 Db 481 GCGCATTTCTGAG 540  
 QY 785 GCTTCGAG 844  
 Db 541 GCTTCGAG 600  
 QY 845 AGCAGAGAGATGCAAAATGACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 904  
 Db 601 AGCAGAGAGATGCAAAATGACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660  
 QY 905 GAAAGTCCGAG 964  
 Db 661 GAAAGTCCGAG 720  
 QY 965 AAGATGATT-GAGAGAGATGCTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1022  
 Db 721 AAGATGATTGAGAGAGATGCTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780  
 QY 1023 AGGAG 1082  
 Db 781 AGAATGAG 840  
 QY 1083 AAG 1116  
 Db 841 GATTAAAAAG 874



RESULT 5  
 BUI51076 846 bp mRNA linear EST 03-SEP-2002  
 LOCUS AGENCOURT\_8586111 Lupski\_symphathetic\_trunk Homo sapiens cDNA clone  
 DEFINITION IMAGE:6195568 5', mRNA sequence.  
 BUI51076  
 ACCESSION BUI51076.1 GI:22664608  
 VERSION EST.  
 KEYWORDS Homo sapiens (human)  
 SOURCE Homo sapiens  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 846)  
 NIH-MGC http://mgs.nci.nih.gov/  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgsabrs@mail.nih.gov  
 Tissue Procurement: Dr. James R. Lupski  
 cDNA Library Preparation: Life Technologies, Inc.  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LLNL13602 row: 3 column: 17  
 High quality sequence stop: 686.  
 Location/Qualifiers  
 1. 846  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:6195568"  
 /sex="male"  
 /tissue\_type="sympathetic trunk"  
 /dev\_stage="adult, 16 yr"  
 /lab\_host="DH10B"  
 /clone\_lib="Lupski\_symphathetic\_trunk"  
 /note="Vector: pCMV-SPORT6 (Life Technologies); Site 1:  
 NotI; Site 2: SalI; cDNA made by oligo-dT priming.  
 Directionally cloned using the following adaptors:  
 5'-TGACCCCAAGCGTCGCG-3' and  
 5'-GACTAGTCTTAGATCGGAGCGCCGCT(15)-3'. Size selected >  
 1 kb for average insert length 1.9 kb. This is a primary  
 library, non-amplified. Library constructed by Life  
 Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor  
 College of Medicine); available through Life  
 Technologies."  
 ORIGIN  
 Query Match 20.8%; Score 768.2; DB 13; Length 846;  
 Query Local Similarity 96.3%; Pred. No. 4.8e-123;  
 Matches 808; Conservative 0; Mismatches 28; Indels 3; Gaps 2;  
 QY 2437 ATGACCCATCAAGAGCAATAGTCTCCAGATTGACGACGACTGAGCAGTATACGAG 2496  
 DB 1 ATGACCCATCAAGAGCAATAGTCTCCAGATTGACGACGACTGAGCAGTATACGAG 60  
 QY 2497 TGCATTTGAGGGAACAAAAGCGCAAAACCTCAAAAGCGGAGCGCTGGATCTTCCTG 2556  
 DB 61 TGCATTTGAGGGAACAAAAGCGCAAAACCTCAAAAGCGGAGCGCTGGATCTTCCTG 120  
 QY 2557 TCTGTGTAAGGTGTACGCAACATCAAGATGTGTGGAGAAAGGGAATGTGTTTCATC 2616  
 DB 121 TCTGTGTAAGGTGTACGCAACATCAAGATGTGTGGAGAAAGGGAATGTGTTTCATC 180  
 QY 2617 CCCCATGACGAGGACACCAATTAAGAAATGCTGGCTTGAAGTAGGGGTTTCTAG 2676  
 DB 181 CCCCATGACGAGGACACCAATTAAGAAATGCTGGCTTGAAGTAGGGGTTTCTAG 240  
 QY 2677 CCCCATCAATGAATGGCTAACTAAACCCCAAGATGAAACCAAGTCACTGCTCCAAACC 2736

DB 241 CCCCATCAATGAATGGCTAACTAAACCCCAAGATGAAACCAAGTCACTGCTCCAAACC 300  
 QY 2737 TTCTGACTTGAGACACGAGACGTATCCAGCAAGCGGAACCTCTGGGAAAACCAATCTGT 2796  
 DB 301 TTCTGACTTGAGACACGAGACGTATCCAGCAAGCGGAACCTCTGGGAAAACCAATCTGT 360  
 QY 2797 GGATTAAGGTCACTTCCCCCACTAAGTTTGAACAGTTCACAGAAAGAACCCAGCTCAAG 2856  
 DB 361 GGATTAAGGTCACTTCCCCCACTAAGTTTGAACAGTTCACAGAAAGAACCCAGCTCAAG 420  
 QY 2857 ACCGAGACACAGCTCACTTGAAGAGGCTAATTCGCTCTGTTTGAATTAATGTAATTT 2916  
 DB 421 ACCGAGACACAGCTCACTTGAAGAGGCTAATTCGCTCTGTTTGAATTAATGTAATTT 480  
 QY 2917 ACTAAATGGGTCATTAATCTTTTATTTTCAATATCCAGTAACCCAGTAATATATTC 2976  
 DB 481 ACTAAATGGGTCATTAATCTTTTATTTTCAATATCCAGTAACCCAGTAATATATTC 540  
 QY 2977 ACTAATTTTATATATCAAGCTAGATGTCATGTAATGTAATGTAATGTAATGTAATTC 3036  
 DB 541 ACTAATTTTATATATCAAGCTAGATGTCATGTAATGTAATGTAATGTAATGTAATTC 600  
 QY 3037 AGCCTGTTTAAAGAAACCATGCTGTGAATAAGACATTTTCTAATGATCATCAATAC 3096  
 DB 601 AGCCTGTTTAAAGAAACCATGCTGTGAATAAGACATTTTCTAATGATCATCAATAC 659  
 QY 3097 TCTGTATCTGAGCGATATACCAACCATGTAAGTCAAGCAAGAGATCCAGTTTAAAT 3156  
 DB 660 TCTGTATCTGAGCGATATACCAACCATGTAAGTCAAGCAAGAGATCCAGTTTAAAT 719  
 QY 3157 TTGCTGCGGAATGTGTGAGTATCTAGAAAATGAACCGTAGTTTGTGTTT- -AA 3214  
 DB 720 TTGCTGCGGAATGTGTGAGTATCTAGAAAATGAACCGTAGTTTGTGTTT- -AA 779  
 QY 3215 ATACAGAGTCATGTTTCTGACATTATTAATTAACATGGAAGAAATTAATCTAGT 3273  
 DB 780 TACAGAAATGATGTTGTTTCTGACATTATTAATTAACATGGAAGAAATTAAT 838  
 RESULT 6  
 CD359921 881 bp mRNA linear EST 29-MAY-2003  
 LOCUS AGENCOURT\_14275922 NIH\_MGC\_180 Homo sapiens cDNA clone  
 DEFINITION IMAGE:30388440 5', mRNA sequence.  
 CD359921  
 ACCESSION CD359921.1 GI:31131332  
 VERSION EST.  
 KEYWORDS Homo sapiens (human)  
 SOURCE Homo sapiens  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 881)  
 NIH-MGC http://mgs.nci.nih.gov/  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 COMMENT Contact: Daniela S. Gerhard, Ph.D.  
 Office of Cancer Genomics  
 National Cancer Institute / NIH  
 Bldg. 31 Rm10A07 Bethesda, MD 20892  
 Email: cgsabrs@mail.nih.gov  
 Tissue Procurement: Dr. Michael Brownstein  
 cDNA Library Preparation: Invitrogen Corp  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: NDAM460 row: 6 column: 01  
 High quality sequence stop: 622.  
 Location/Qualifiers  
 1. 881  
 /organism="Homo sapiens"

## ORIGIN

/mol\_type="rRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:30388440"  
 /lab\_host="DH10B-Ton A (T1 and T5 phage resistancees)"  
 /clone\_lib="NIH MGC 180"  
 /note="Organ: Testis; Vector: pCMV-Sport6.1; Site 1: NotI;  
 Site 2: EcoRV (destroyed); Library is oligo-dT primed and  
 directionally cloned (EcoRV site is destroyed upon  
 cloning). Average insert size 1.68 kb. Library was  
 constructed by (Invitrogen). Note: this is a NIH\_MGC  
 library."

Query Match 20.7%; Score 768; DB 14; Length 881;  
 Best Local Similarity 99.2%; Pred. No. 5.1e-123;  
 Matches 771; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 CCTAGGGAATGACAGGCACTCTCCACAGGAGGCTGCATCCACTTGGCTGGCTGTC 61  
 1 CCTAGGGAATGACAGGCACTCTCCACAGGAGGCTGCATCCACTTGGCTGGCTGTC 60  
 QY 62 ATTGGCTGCTTATGAAAAACGACAGGCAATGCAATACACCGCTCCCGACTGTAAC 121  
 61 ATTGGCTGCTTATGAAAAACGACAGGCAATGCAATACACCGCTCCCGACTGTAAC 120  
 QY 122 ATAGGGGATATGTGTCACTTACATGATGACTTCTGGAGGGGCAAGAAAGGCGTCTG 181  
 121 ATAGGGGATATGTGTCACTTACATGATGACTTCTGGAGGGGCAAGAAAGGCGTCTG 180  
 QY 182 GAGTTTATTAATGATGACAGAGTGTATTCGGCTGCTGCGCCGCTGCTGCTC 241  
 181 GAGTTTATTAATGATGACAGAGTGTATTCGGCTGCTGCGCCGCTGCTGCTC 240  
 QY 242 TGGCTGTGCTCTGTAAAGAAATCAAGTCTCTTCCTTCGACTTATGCTCGGAGAA 301  
 241 TGGCTGTGCTCTGTAAAGAAATCAAGTCTCTTCCTTCGACTTATGCTCGGAGAA 300  
 QY 302 GTTTCAGCTCAAGGTATCATTTGAAACATTTCAAGATCATCAATCAATTCACAGGG 361  
 301 GTTTCAGCTCAAGGTATCATTTGAAACATTTCAAGATCATCAATCAATTCACAGGG 360  
 QY 362 ATTGGTGAACCAAGAGGCTGACATCTGATTTGCTGACTGCTTCACACATCATCTGG 421  
 361 ATTGGTGAACCAAGAGGCTGACATCTGATTTGCTGACTGCTTCACACATCATCTGG 420  
 QY 422 TCTCCCTGAACCTGAATCAACATGATGATTTGAGGCTGCGAGAGACTTGAAGG 481  
 421 TCTCCCTGAACCTGAATCAACATGATGATTTGAGGCTGCGAGAGACTTGAAGG 480  
 QY 482 CAAAAGAGGAGAGATGCGACTCGAAGCAAAAGAAATGGCTTACCAAGAAATGACAT 541  
 481 CAAAAGAGGAGAGATGCGACTCGAAGCAAAAGAAATGGCTTACCAAGAAATGACAT 540  
 QY 542 GATGAG 601  
 541 GATGAG 600  
 QY 602 AAGCAG 661  
 601 AAGCAG 660  
 QY 662 AGTGTGCTGACGAT 721  
 661 AGTGTGCTGACGAT 720  
 QY 722 GAGGCGGCAATCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 778  
 721 GAGGCGGCAATCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 777

RESULT 7  
 CD643635 827 bp mRNA linear EST 17-JUN-2003  
 LOCUS CD643635

DEFINITION AGENCOURT\_14553720 NIH Human H1 Embryonic Stem Cell cDNA library  
 (long) Homo sapiens cDNA clone IMAGE:30425920 5', mRNA sequence.  
 ACCESSION CD643635  
 VERSION CD643635.1 GI:31814712  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 REFERENCE 1 (bases 1 to 827)  
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Daniela S. Gerhard, Ph.D.  
 Office of Cancer Genomics / NIH  
 National Cancer Institute / NIH  
 Bldg. 31 Rm10A07 Bethesda, MD 20892  
 Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
 Tissue Procurement: Irene Gins and Mahendra Rao, NIA  
 cDNA Library Preparation: Yulan Piao and Minoru Ko  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 clone distribution: MGC c  
 can be found through the I.M.A.G.E. Consortium/LNL at:  
<http://image.llnl.gov>  
 plate: NDM510 row: 0 column: 17  
 High quality sequence stop: 662.  
 Location/Qualifiers

FEATURES  
 source

1..827  
 /organism="Homo sapiens"  
 /mol\_type="rRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:30425920"  
 /tissue\_type="Embryonic Stem cells"  
 /cell\_line="WA01"  
 /lab\_host="DH10B (T1 phage-resistant)"  
 /clone\_lib="NIA Human H1 Embryonic Stem Cell cDNA library (long)"  
 /note="Vector: pCMV-Sport6; Site 1: NotI; Site 2: SalI;  
 This is a long-transcript enriched cDNA library (genome  
 Res. 11: 1553-1558 (2001). [PMID: 11544199]) from WA01  
 cell line. Undifferentiated human ES cell line WA01/H1  
 was obtained from Wicell Research Institute, Inc.  
 Madison, WI, cultured according to their instructions, on  
 MEf feeders. They formed round colonies with defined edges  
 and were positive for alkaline phosphatase, SSEA-4, OCT3,  
 OCT4, REX1, UTR, TERT, SOX2, CX43 and CX45. They are  
 negative for GATA2, GATA4, PDX1, NCAM, MSX1, FLK3, SSEA-1,  
 TUBB3, NES, GFAP, and BOMBS. When confluent (18-10 days  
 after plating), the ES cells from 4 X 6cm dishes were  
 treated with 1 mg/ml collagenase, type IV  
 (Invitrogen/Gibco) for 5-10 min and gently scraped off  
 with 5 ml pipette. RNA was purified with RNeasy Reagent  
 from Invitrogen. Protocol ref: Genome Res. 11: 1553-1558  
 (2001). [PMID: 11544199] Double-stranded cDNAs were  
 synthesized with an oligo(dT) primer [Invitrogen:  
 5'-pGACTAGTCTGATCGCAGCGCGCCCTTTTCTTTTCTTTT-3'] from  
 3.4g of total RNA, treated with T4 DNA polymerase, and  
 purified by ethanol-precipitation. The cDNAs were ligated  
 to Lene-linker LB-SalI, purified by phenol/chloroform  
 extraction, and separated from free linkers by  
 Centricon-100 column. Then, the cDNAs were amplified by  
 long-range high fidelity PCR using Ex tag polymerase  
 (Takara) with a primer Sal4-S for 25 cycles. The products  
 were purified by phenol/chloroform extraction and  
 Centricon-100 column. The cDNAs were digested with SalI  
 and NotI enzymes and cloned into SalI/NotI site of  
 pCMV-Sport6 plasmid vector. The average insert size is  
 about 3.6kb.

## ORIGIN

Query Match 20.4%; Score 753.8; DB 14; Length 827;  
 Best Local Similarity 97.5%; Pred. No. 1.5e-120;



QY 801 AGGAGTTCGACCCACATATAAGATGCAAGTCTGTCTGCCAAGCAGAGATGCAAA 860  
 Db 601 AGGAGTTCGACCCACATATAAGATGCAAGTCTGTCTGCCAAGCAGAGATGCAAA 660  
 QY 861 ATGACACACGAGAAATTAATCTACCGGAGAGAGAAATTAATCTGCCAGAGAA 920  
 Db 661 ATGACACACGAGAAATTAATCTACCGGAGAGAGAAATTAATCTGCCAGAGAA 720  
 QY 921 GATACGAGATGAGAGAAACAGAAACAGTC 949  
 Db 721 GATACGAGATGAGAGAAACAGAAACAGTC 749

RESULT 9  
 AUI32922 812 bp mRNA linear EST 01-AUG-2002  
 LOCUS AUI32922 NT2RP4 Homo sapiens cDNA clone NT2RP4000885 5', mRNA  
 DEFINITION  
 AUI32922.1 GI:10993461  
 VERSION AUI32922.1  
 KEYWORDS Homo sapiens (human)  
 SOURCE Homo sapiens  
 ORGANISM Homo sapiens  
 REFERENCE Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 1 (bases 1 to 812)  
 AUTHORS Ota, T., Sugiyama, T., Ishii, S., Suzuki, Y., Saito, K., Yamamoto, J., Nishikawa, T., Nakamura, Y., Nagai, T., Sugano, S., Masuko, Y., and Isegai, T.  
 TITLE HRI human cDNA project (Ota, T., Sugiyama, T., Ishii, S., Suzuki, Y., Saito, K., Yamamoto, J., Nishikawa, T., Nakamura, Y., Nagai, T., Sugano, S., Masuko, Y., Isegai, T.)  
 JOURNAL Unpublished (2000)  
 COMMENT Contact: Takao Isegai  
 Genomics Laboratory  
 Helix Research Institute  
 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan  
 Tel: 81-438-52-3975  
 Fax: 81-438-52-3986  
 Email: genomics@hri.co.jp  
 HRI human cDNA project; 5'- & 3'-end one pass sequencing; Helix Research Institute; cDNA library construction; Department of Virology, Institute of Medical Science, University of Tokyo, and Helix Research Institute.  
 FEATURES  
 source  
 1. 812  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="NT2RP4000885"  
 /cell\_type="teratocarcinoma"  
 /cell\_line="NT2"  
 /clone\_1b="NT2RP4"  
 /note="Vector: pME18SFL3; mRNA from NT2 neuronal precursor cells after 2-weeks retinoic acid (RA) induction"

Query Match 19.9%; Score 737.8; DB 9; Length 812;  
 Best Local Similarity 97.9%; Fred. No. 8.9e-118;  
 Matches 799; Conservative 0; Mismatches 10; Indels 7; Gaps 5;

QY 1910 ACAGAAAGTTAAGTCGAGATGAGATTCATGACCCACAACTTAACATCTAGAGAT 1969  
 Db 1 ACAGAAAGTTAAGTCGAGATGAGATTCATGACCCACAACTTAACATCTAGAGAT 60

QY 1970 ACTTTTCAGCCCGCTGAGAGAGAGGCGCAGGCTGAGACCAAGAGGCTGAGGCGCC 2029  
 Db 61 ACTTTTCAGCCCGCTGAGAGAGAGGCGCAGGCTGAGACCAAGAGGCTGAGGCGCC 120

QY 2030 CAGGTGAAGCCGCGCAAAAGGCTGAGAGCTTGTGTGTGTGTGTGTGTGTGTGTGT 2089  
 Db 121 CAGGTGAAGCCGCGCAAAAGGCTGAGAGCTTGTGTGTGTGTGTGTGTGTGTGTGT 180

QY 2090 GAAGATTCGAGAAAGCTCAACAGAGAGCGCAGAGCGGCTTTGGAGCTGAGAGAACTC 2145  
 Db 181 GAAGATTCGAGAAAGCTCAACAGAGAGCGCAGAGCGGCTTTGGAGCTGAGAGAACTC 240

QY 2150 AAGAAAAAGAGGAG 2209  
 Db 241 AAGAAAAAGAGGAG 300

QY 2210 GAGGAGCGGATTCGAAAACTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2269  
 Db 301 GAGGAGCGGATTCGAAAACTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360

QY 2270 GAAAGCGGAG 2329  
 Db 361 GAAAGCGGAG 420

QY 2330 GATGACAAAGAAACCATTCAGAGTGTCTTCACTCTTAAAGTTTCAATCTCAAGATG 2389  
 Db 421 GATGACAAAGAAACCATTCAGAGTGTCTTCACTCTTAAAGTTTCAATCTCAAGATG 480

QY 2390 CGAGCAGAAATTTTGAATPAAGTCTGTGCAAGAAAGAGAGAGAGAGAGAGAGAG 2449  
 Db 481 CGAGCAGAAATTTTGAATPAAGTCTGTGCAAGAAAGAGAGAGAGAGAGAGAGAG 537

QY 2450 GCAGCAATATGCTCCAGATTGACAGACAGACTGAGACAGATACCAAGTGCATTTGAG 2509  
 Db 538 GCAGCAATATGCTCCAGATTGACAGACAGACTGAGACAGATACCAAGTGCATTTGAG 597

QY 2510 ACAAAGACCGCAAAACCTTCAAG--CGGACGCTGGAGTCTTCTGCTGCGAAG 2568  
 Db 598 ACAAAGACCGCAAAACCTTCAAG--CGGACGCTGGAGTCTTCTGCTGCGAAG 657

QY 2569 TGTACGCAACATCAAGATATGTGGAGAAAGGAAATGTGTTTCATCTCCAC--TGCAG 2627  
 Db 658 TGTACGCAACATCAAGATATGTGGAGAAAGGAAATGTGTTTCATCTCCAC--TGCAG 717

QY 2628 CAGGACACCAATTAAGAAAC--TGCTGCTTGAAGTGAAGGTTTGAAGCGATCAAT 2686  
 Db 718 CAGGACACCAATTAAGAAAC--TGCTGCTTGAAGTGAAGGTTTGAAGCGATCAAT 776

QY 2687 GATGCTTAATTAACCCAGATGAGAAACCACTCA 2722  
 Db 777 GATGCTTAATTAACCCAGATGAGAAACCACTCA 812

RESULT 10  
 CD656598 823 bp mRNA linear EST 18-JUN-2003  
 LOCUS AGENCOURT 14540446 NIA Human HI Embryonic Stem Cell cDNA Library  
 DEFINITION (Long) Homo sapiens cDNA clone IMAGE:30419432 5', mRNA sequence.  
 CD656598  
 ACCESSION CD656598.1 GI:31898079  
 VERSION EST.  
 KEYWORDS Homo sapiens (human)  
 SOURCE Homo sapiens  
 ORGANISM Homo sapiens  
 REFERENCE Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 1 (bases 1 to 823)  
 AUTHORS NIH-MGC http://img.nci.nih.gov/  
 TITLE NIH-MGC  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Daniela S. Gerhard, Ph.D.  
 Office of Cancer Genomics  
 National Cancer Institute / NIH  
 3149 31 Rm10A07 Bethesda, MD 20892  
 Email: cgsabbs-remail.nih.gov  
 Tissue Procurement: Irene Givis and Mahendra Rao, NIA  
 cDNA Library Preparation: Yulan Piao and Minoru Ko  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNU)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC  
 can be found through the I.M.A.G.E. Consortium/ILNU at:  
 http://image.llnl.gov

Plate: NDAM494 row: a column: 09  
High quality sequence stop: 695.

## FEATURES

source

1. .823

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:30419432"

/tissue\_type="Embryonic Stem cells"

/cell\_line="WA01"

/lab\_host="DH10B (T1 phage-resistant)"

/clone\_lib="NIA Human H1 Embryonic Stem Cell cDNA Library (long)"

/note="Vector: pCMV-Sport6, Site 1: NotI, Site 2: SalI, this is a long-transcript enriched cDNA library (Genome Res. 11: 1553-1558 (2001)). [PMID: 11541991] from WA01 cell line. Undifferentiated human ES cell line WA01/H1 was obtained from WiCell Research Institute, Inc., Madison, WI, cultured according to their instructions, on MEK feeders. They formed round colonies with defined edges and were positive for alkaline phosphatase, SSEA-4, OCT3, OCT4, REX1, UTR, TERT, SOX2, CX43 and CX45. They are negative for GAP2, GAP24, PDX1, NCAM, MSX1, FLT3, SSEA-1, TUBB3, NES, GFAP, and BOMES. When confluent (18-10 days after plating), the ES cells from 4 X 6cm dishes were treated with 1 mg/ml collagenase, type IV (Invitrogen/GIBCO) for 5-10 min and gently scraped off with 5 ml pipette. RNA was purified with TRIzol Reagent from Invitrogen. Protocol ref: Genome Res. 11: 1553-1558 (2001). [PMID:11541991] Double-stranded cDNAs were synthesized with an Oligo(dT) primer (Invitrogen: 5'-pGAGTGTCTAGATCGGAGCGCGCCCTTTTCTTTT-3') from 3.4g of total RNA, treated with T4 DNA polymerase, and purified by ethanol-precipitation. The cDNAs were ligated to Loner-linker IL-Sal4, purified by phenol/chloroform extraction, and separated from free linkers by Centricon-100 column. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Taq polymerase (Takara) with a primer Sal4-S for 25 cycles. The products were purified by phenol/chloroform extraction and Centricon-100 column. The cDNAs were digested with SalI and NotI enzymes and cloned into SalI/NotI site of pCMV-Sport6 plasmid vector. The average insert size is about 3.6kb."

## ORIGIN

Query Match 19.8%; Score 732; DB 14; Length 823;  
Best Local Similarity 97.5%; Pred. No. 9e-117;  
Matches 752; Conservative 0; Mismatches 18; Indels 1; Gaps 1;

QY 260 AAGAAATCAGTCTCTCTTCCGACTTATGCTCGGGAAGAAAGTTTCAAGTCAAGTCA 319  
Db 12 AAGAAATCAGTCTCTCTTCCGACTTATGCTCGGGAAGAAAGTTTCAAGTCAAGTCA 71  
QY 320 TCATTGGAACATTTCAAGATCATCAATCAATTCACAGGATTTGTGACCAACAGAA 379  
Db 72 TCATTGGAACATTTCAAGATCATCAATCAATTCACAGGATTTGTGACCAACAGAA 131  
QY 380 GGGTCAGACATCTGATTGCTGACCTGTCCAGACATCATCTGCTCTCCCTGAACCTGAAT 439  
Db 132 GGGTCAGACATCTGATTGCTGACCTGTCCAGACATCATCTGCTCTCCCTGAACCTGAAT 191  
QY 440 CACACATGATGATTTTGAAGCGTCGACAGAACTTAAAGGCAAAAGAGAGAGATG 499  
Db 192 CACACATGATGATTTTGAAGCGTCGACAGAACTTAAAGGCAAAAGAGAGAGATG 251  
QY 500 CGACTGGAAGCAAGAAAGATGCTCTACAGAGAAATGACATGATGAAAGAGAGAGAGCC 559  
Db 252 CGACCCGGAACAGAAAGATGCTCTACAGAGAAATGACATGATGAAAGAGAGAGAGCC 311  
QY 560 CGGAAAGCGGCGCGCGAGCGGACAGAGACGGCTGCGGCGAGAAACAGAGAGAGATCC 619  
Db 312 CGGAAAGCGGCGCGCGAGCGGACAGAGACGGCTGCGGCGAGAAACAGAGAGAGATCC 371

QY 620 TTGGGACAGGTGACCGACACAGGTGAGGTGATGCCCAGAAACAGTGTGCTGACGAGAG 679  
Db 372 TTGGGACAGGTGACCGACACAGGTGAGGTGATGCCCAGAAACAGTGTGCTGACGAGAG 431  
QY 680 GCCAAGCAACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 729  
Db 432 GCCAAGCAACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 491  
QY 740 CGCTGCTCGGCGGTGAGAAAGACGCCCAAAACGCCCTTCAGAGGCTTGAAGCGGAG 799  
Db 492 CGCTGCTCGGCGGTGAGAAAGACGCCCAAAACGCCCTTCAGAGGCTTGAAGCGGAG 551  
QY 800 AAGAGTTCAACCCCAACCAATTAACAGATGCAAGTCTGTGCTCCCAAGCAAGAAATGCA 859  
Db 552 AAGAGTTCAACCCCAACCAATTAACAGATGCAAGTCTGTGCTCCCAAGCAAGAAATGCA 611  
QY 860 AATGACACAGCAGAAATGAACTACCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 919  
Db 612 AATGACACAGCAGAAATGAACTACCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 670  
QY 920 AGATGACAGTAAAG 979  
Db 671 AGATGACAGTAAAG 730  
QY 980 GATGCTGAAGAAACAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1030  
Db 731 GATGCTGAAGAAACAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 781

RESULT 11  
LOCUS BC010124/c  
DEFINITION UT-H-ED0-axz-o-23-0-UT-s1 NCI CGAP\_ED0 Homo sapiens cDNA clone  
IMAGE:5835886 3', mRNA sequence.

ACCESSION BC010124  
VERSION BC010124.1  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE NCI CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
TITLE Tumor Gene Index

JOURNAL Unpublished (1997)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)

Tissue Procurement: Dr. Jose Mercende  
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Clone distribution information can be found  
through the I.M.A.G.E. Consortium/INM at: <http://image.llnl.gov>  
The following repetitive elements were found in this cDNA  
sequence: 1-21, >At rich#low\_complexity (matched complement)  
Seq primer: M13 FORWARD  
PolyAyes.

FEATURES  
Location/Qualifiers  
1. .748

/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5835886"  
/tissue\_type="Chondrosarcoma"  
/dev\_stage="adult"  
/lab\_host="DH10B (Life Technologies)"  
/clone\_id="NCI CGAP\_ED0"  
/note="Organ: Left Femur Bone, Vector: pTTT3-Pac  
(Pharmacia) with a modified polylinker, Site 1: EcoR I,  
Site 2: Not I; NCI CGAP ED0 is a cDNA library containing  
the following tissue(s): Chondrosarcoma cell line C85. The

library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT733-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dri)18 tail. The sequence tag for this library is GCTCAAGGCT.  
TAG LIB=UI-H-ED0  
TAG\_SEQ=CGTCAAGGCT"

## ORIGIN

Query Match 19.6%; Score 725.4; DB 12; Length 748;  
Best Local Similarity 99.1%; Pred. No. 1.3e-115;  
Matches 740; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

2611 TTCAATCCCACTGACGAGGACACCAATTAAGAACTGCTGCTGAAGTAGGGGT 2670  
748 TTCAATCCCACTGACGAGGACACCAATTAAGAACTGCTGCTGAAGTAGGGGT 690  
2671 TTCAAGCGCATGCAATGATGCTAACTAAACCCAGATGAGAAACAGTCACTGCTCC 2730  
689 TTCAAGCGCATGCAATGATGCTAACTAAACCCAGATGAGAAACAGTCACTGCTCC 630  
2731 CAAACCTTCTGACTTGAGACCAAGAGAGCTATCCAGAAACCGAATCTTGAGAAAGCA 2790  
629 CAAACCTTCTGACTTGAGACCAAGAGAGCTATCCAGAAACCGAATCTTGAGAAAGCA 570  
2791 ATCTGTGATAGGACCTTCCCACTTAAGTTGAGACAGTTCCAGAAAGAACCAAG 2850  
569 ATCTGTGATAGGACCTTCCCACTTAAGTTGAGACAGTTCCAGAAAGAACCAAG 510  
2851 CTCAAGACGACGAGAGAGTCAAGTTGAGAGGCTTAATGCTGCTGTTTATTTATGT 2910  
509 CTCAAGACGACGAGAGAGTCAAGTTGAGAGGCTTAATGCTGCTGTTTATTTATGT 450  
2911 TGATTTACTAATGGGTTCAATATCTTTATTTTCAATATCCAGTAACCCATGAT 2970  
449 TGATTTACTAATGGGTTCAATATCTTTATTTTCAATATCCAGTAACCCATGAT 390  
2971 ATTATCACTAATTTAATTAATCAAGTCTAGAGTCTCAATGCTTAAGTCACTGCTTG 3030  
389 ATTATCACTAATTTAATTAATCAAGTCTAGAGTCTCAATGCTTAAGTCACTGCTTG 330  
3031 CACAGGAGCCTGTTCTAAGAAACCCATGCTGTGAATAGAGACTTTTCTACTGATCAT 3090  
329 CACAGGAGCCTGTTCTAAGAAACCCATGCTGTGAATAGAGACTTTTCTACTGATCAT 270  
3091 CATATCTGTATCTGAGAGAGTATCCAAACCACTGTAAGTCAACAGAGATCCAGT 3150  
269 CATATCTGTATCTGAGAGAGTATCCAAACCACTGTAAGTCAACAGAGATCCAGT 210  
3151 TTAATATGCTGCGGAATGTGTGAGATCTAGAAATGAACCGTAGTTTGTGTTT 3210  
209 TTAATATGCTGCGGAATGTGTGAGATCTAGAAATGAACCGTAGTTTGTGTTT 150  
3211 TTAATATGAGAGTATGTTGTTTCTGCACTTATATTAAGATGAGAAATATATCTT 3270  
149 TTAATATGAGAGTATGTTGTTTCTGCACTTATATTAAGATGAGAAATATATCTT 90  
3271 AGTAGGCAATGTAACTTTTGAAGTAACCATTTTCAAGTTTGAAGTAACCATTTT 3330  
89 AGTAGGCAATGTAACTTTTGAAGTAACCATTTTCAAGTTTGAAGTAACCATTTT 30  
3331 TGGTGTCTTAAAAAAGAAAAAGA 3357  
29 TGGTGTCTTAAAAAAGAAAAAGA 3

RESULT 12

AI798679/c  
LOCUS AI798679 775 bp mRNA linear EST 06-JUL-1999  
DEFINITION w91c12.X1 Soares\_NFL\_T\_GBC\_S1 Homo sapiens cDNA clone IMAGE:2348470 3' similar to SW:CALD\_HUMAN Q05682 CALDESMON ; mRNA sequence.

ACCESSION AI798679  
VERSION AI798679  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 775)  
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index

JOURNAL Unpublished (1997)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapb-remail.nih.gov

THIS clone is available royalty-free through LNL; contact the IMAGE Consortium ([info@image.llnl.gov](mailto:info@image.llnl.gov)) for further information.  
Seq primer: -40UP from Gibco  
High quality sequence stop: 465.  
Location/Qualifiers

1..775  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:2348470"  
/lab\_host="DH10B"  
/clone\_lib="Soares\_NFL\_T\_GBC\_S1"

/note="Organ: pooled; Vector: pT733-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; Equal amounts of plasmid DNA from three normalized libraries (fetal lung NbH19W, testis NHT, and B-cell NCI CGAP GCBI) were mixed, and as circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pool consisted of 1.M.A.G.E. clones 297480-302087, 682632-687239, 726408-728711, and 729096-733399. Subtraction by Bento Soares and M. Fatima Bonaldo."

## ORIGIN

Query Match 19.4%; Score 719.6; DB 9; Length 775;  
Best Local Similarity 97.5%; Pred. No. 1.3e-114;  
Matches 750; Conservative 0; Mismatches 17; Indels 2; Gaps 2;

2573 CGCAACATCAAGATATGTGGAGAAAGGAATGTGTTTCATCCCACTGACGAGGC 2632  
769 CGCAACATCAAGATATGTGGAGAAAGGAATGTGTTTCATCCCACTGACGAGGC 711  
2633 ACACC-AAATAGAGAACTGCTGCTTGAAGTAGGGGTTTCTAGCCGATCAATGATG 2691  
710 ACACC-AAATAGAGAACTGCTGCTTGAAGTAGGGGTTTCTAGCCGATCAATGATG 651  
2692 GCTAATCAAAACCCAGATGAAACAGTCACTGCTCCAAACCTTCTGACTGAGACC 2751  
650 NCTAATCTAGAGCCAGAAAGAAACAGTCACTGCTCCAAACCTTCTGACTGAGACC 591  
2752 AGGAGAGTATCAGAGACGGAACCTCTGGGAAAGCAATCTGTGTATAGTCACTTC 2811  
590 AGGAGAGTATCAGAGACGGAACCTCTGGGAAAGCAATCTGTGTATAGTCACTTC 531  
2812 CCCCACTAAGTTTGAAGATCCAGAAAGAACCAAGTCAAGAGGAGAGAGAGTTC 2871  
530 CCNCACTAAGTTTGAAGATCCAGAAAGAACCAAGTCAAGAGGAGAGAGAGTTC 471  
2872 AGTTGTAGAGGCTAATGCTGCTGTTTGTATTTATGTTGATTTACTAAATGGCTTCA 2931  
470 AGTTGTAGAGGCTAATGCTGCTGTTTGTATTTATGTTGATTTACTAAATGGCTTCA 411



QY 2932 TTAATCTTTATTTTTCATATCCAGTAACCCATGTAATATATATATATATAT 2991  
 DB 410 TTAATCTTTATTTTTCATATCCAGTAACCCATGTAATATATATATATAT 351  
 QY 2992 CACAGCTAGAGATGCTTCATGCTAAAGTACGCTTTCACAGAGAGCTGTTCTTAAG 3051  
 DB 350 CACAGCTAGAGATGCTTCATGCTAAAGTACGCTTTCACAGAGAGCTGTTCTTAAG 291  
 QY 3052 AAACCCATGCTGTAAATAGAGACTTTTCTACTATCATATATCTTGTATCTGAGCAG 3111  
 DB 290 AAACCCATGCTGTAAATAGAGACTTTTCTACTATCATATATCTTGTATCTGAGCAG 231  
 QY 3112 TGATACCAACACATCTGAAAGTCAACAGAGATCCAGATTTAAATGCTGCGAGATGT 3171  
 DB 230 TGATACCAACACATCTGAAAGTCAACAGAGATCCAGATTTAAATGCTGCGAGATGT 171  
 QY 3172 GTGCAGATGCTAGAAAATGAAACCGTAGTTTGTGTTTAAATACAGAGATGCTG 3231  
 DB 170 GTGCAGATGCTAGAAAATGAAACCGTAGTTTGTGTTTAAATACAGAGATGCTG 111  
 QY 3232 TTTCTGACCTTATATATTAAGCATGAAAGAAATTAATCTTAGAGCAATTTGTAACACTTT 3291  
 DB 110 TTTCTGACCTTATATATTAAGCATGAAAGAAATTAATCTTAGAGCAATTTGTAACACTTT 51  
 QY 3292 TTGAAGTAACCCATTCAGATTTGAATATCTGCGATATATGTTGCTTT 3340  
 DB 50 TTGAAGTAACCCATTCAGATTTGAATATCTGCGATATATGTTGCTTT 2

RESULT 13  
 LOCUS A1926727/c 753 bp mRNA linear EST 08-MAR-2000  
 DEFINITION A1926727.1 NC1 CGAP Gas4 Homo sapiens cDNA clone IMAGE:2458900.3,  
 similar to SW:GALD HUMAN 005682 CALDESMON; contains element TAR1  
 repetitive element; mRNA sequence.

ACCESSION A1926727  
 VERSION A1926727.1 GI:5662691  
 KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE 1 (bases 1 to 753)  
 NC1-GCAP http://www.ncbi.nlm.nih.gov/nc1gcgcp.

AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index  
 JOURNAL Unpublished (1997)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgaps-remail.nih.gov  
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.  
 Emmert-Buck, M.D., Ph.D.  
 CDNA Library Preparation: Life Technologies, Inc.  
 DNA Sequencing: Arrayed by: Greg Lennon, Ph.D.  
 Cloning Distribution: NCI-GCAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/BLN at:  
 www-bio.lit.nih.gov/bdtp/image/image.html  
 Insert Length: 3783 Std Error: 0.00  
 Seq primer: -40UP from Gibco  
 High quality sequence stop: 396.

## FEATURES

Location/Qualifiers  
 1..753  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /cd\_xref="taxon:9606"  
 /clone="IMAGE:2458900"  
 /tissue\_type="poorly differentiated adenocarcinoma with  
 signet ring cell features"  
 /lab\_host="DH10B"  
 /clone\_id="NC1 CGAP Gas4"  
 /note="Organ: stomach; Vector: pCMV-SPORT6; Site: 1; Salt:  
 Site 2: NCI; Cloned unidirectionally. Primer: Oligo dT.  
 Average insert size 1.69 kb. Life Technologies catalog #:

ORIGIN 11549-011"

Query Match 19.1%; Score 708.8; DB 9; Length 753;  
 Best Local Similarity 96.3%; Pred. No. 9.6e-113;  
 Matches 725; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

QY 1082 GAAGAGAAAACAACCTGAAGCCAGAGAGAAACAGTGTATGCTATTAATAAATGGGCGAG 1141  
 DB 753 GAGCAGCCAGCAGCTTGAGAGCCAGCAGGCAACAGTGTATGCTATTAACAATGGGCGAG 694  
 QY 1142 ATCAGTTTCAGAAAGCCTTAAACAAAGAGAGAGAGGAAACAAGTTCAATGAGATTTC 1201  
 DB 693 ATCAGTCAAGAAAGCCTTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATTTC 634  
 QY 1202 CATCATGAAAG 1261  
 DB 633 CATCATGAAAG 574  
 QY 1262 GAAGCAG 1321  
 DB 573 GAAGCAG 514  
 QY 1322 GCAAGAAATTGAAGCAG 1381  
 DB 513 GCAAGAAATTGAAGCAG 454  
 QY 1382 GAGAGGAG 1441  
 DB 453 GAGAGGAG 394  
 QY 1442 GAG 1501  
 DB 393 GAG 334  
 QY 1502 GCAAG 1561  
 DB 333 GCAAG 274  
 QY 1562 GCCAGGAG 1621  
 DB 273 GCCAGGAG 214  
 QY 1622 GAAG 1681  
 DB 213 GAAG 154  
 QY 1682 ATAG 1741  
 DB 153 ATAG 94  
 QY 1742 CTAAG 1801  
 DB 93 CTAAG 34  
 QY 1802 GAAG 1834  
 DB 33 GAAG 1  
 RESULT 14  
 LOCUS AUI40904 799 bp mRNA linear EST 05-AUG-2002  
 DEFINITION AUI40904 PLACE4 Homo sapiens cDNA clone PLACE4000466.5', mRNA  
 sequence.  
 ACCESSION AUI40904  
 VERSION AUI40904.1 GI:11002425  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 REFERENCE 1 (bases 1 to 799)

**AUTHORS**  
Ota, T., Suzuki, Y., Saito, K., Iehi, S., Yamamoto, J., Sugiyama, T.,  
Nishikawa, T., Nakamura, Y., Sugano, S., Maehno, Y., and Isega, T.  
HRI human cDNA project (Ota, T., Suzuki, Y., Saito, K., Iehi, S.,  
Yamamoto, J., Sugiyama, T., Nishikawa, T., Nakamura, Y., Sugano, S.,  
Maehno, Y., Isega, T.)  
Unpublished (2000)  
**JOURNAL**  
Contact: Takao Isega  
Genomics Laboratory  
Helix Research Institute  
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan  
Tel: 81-438-52-3975  
Fax: 81-438-52-3986  
Email: genomics@hri.co.jp  
HRI human cDNA project, 5'- & 3'-end one pass sequencing: Helix  
Research Institute; cDNA library construction: Department of  
Virology, Institute of Medical Science, University of Tokyo, and  
Helix Research Institute.  
**FEATURES**  
Location/Qualifiers  
1..799  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="PLACE4000466"  
/tissue\_type="Placenta"  
/clone\_lib="PLACE4"  
/note="Vector: pME18SFL3"  
**ORIGIN**  
Query Match 19.1%; Score 708.8; DB 9; Length 799;  
Best Local Similarity 99.4%; Pred. No. 9.5e-113; Indels 1; Gaps 1;  
Matches 721; Conservative 0; Mismatches 3;  
490 GGAGGAGATGCGACTGGAAGCAAGAAATGCGCTACAGAGAAATGATGATGAAGA 549  
20 GGAGGAGATGCGACTGGAAGCAAGAAATGCGCTACAGAGAAATGATGATGAAGA 79.  
550 GGAGGAGATGCGACTGGAAGCAAGAAATGCGCTACAGAGAAATGATGATGAAGA 609  
80 GGAGGAGATGCGACTGGAAGCAAGAAATGCGCTACAGAGAAATGATGATGAAGA 139  
610 GGAGGAGATGCGACTGGAAGCAAGAAATGCGCTACAGAGAAATGATGATGAAGA 669  
140 GGAGGAGATGCGACTGGAAGCAAGAAATGCGCTACAGAGAAATGATGATGAAGA 199  
670 GGAGGAGATGCGACTGGAAGCAAGAAATGCGCTACAGAGAAATGATGATGAAGA 729  
200 GGAGGAGATGCGACTGGAAGCAAGAAATGCGCTACAGAGAAATGATGATGAAGA 259  
730 GGAGGAGATGCGACTGGAAGCAAGAAATGCGCTACAGAGAAATGATGATGAAGA 789  
260 GGAGGAGATGCGACTGGAAGCAAGAAATGCGCTACAGAGAAATGATGATGAAGA 319  
790 GGAGGAGATGCGACTGGAAGCAAGAAATGCGCTACAGAGAAATGATGATGAAGA 849  
330 GGAGGAGATGCGACTGGAAGCAAGAAATGCGCTACAGAGAAATGATGATGAAGA 379  
850 GGAGGAGATGCGACTGGAAGCAAGAAATGCGCTACAGAGAAATGATGATGAAGA 909  
380 GGAGGAGATGCGACTGGAAGCAAGAAATGCGCTACAGAGAAATGATGATGAAGA 439  
910 GGAGGAGATGCGACTGGAAGCAAGAAATGCGCTACAGAGAAATGATGATGAAGA 969  
440 GGAGGAGATGCGACTGGAAGCAAGAAATGCGCTACAGAGAAATGATGATGAAGA 499  
970 GGAGGAGATGCGACTGGAAGCAAGAAATGCGCTACAGAGAAATGATGATGAAGA 1029  
500 GGAGGAGATGCGACTGGAAGCAAGAAATGCGCTACAGAGAAATGATGATGAAGA 559  
1030 GGAGGAGATGCGACTGGAAGCAAGAAATGCGCTACAGAGAAATGATGATGAAGA 1089  
560 GGAGGAGATGCGACTGGAAGCAAGAAATGCGCTACAGAGAAATGATGATGAAGA 619  
1090 AACCACTGAAGCGAGAGAAACAGTGTATGTCTTAAAAAATGGCAGATCAGTTC 1149

**RESULT 15**  
BX642011 735 bp mRNA linear EST 04-SEP-2003  
LOCUS DKFZP686H16201.1 686 (synonym: hicc3) Homo sapiens cDNA clone  
DEFINITION DKFZP686H16201.1 GI:34476330  
ACCESSION BX642011  
VERSION BX642011.1 GI:34476330  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE 1 (bases 1 to 735)  
Poustka, A., Albert, R., Moosmayer, P., Schupp, I., Wellenreuther, R.,  
Mewes, R.W., Well, B., Amd, C., Oesinger, A., Fob, G., Han, W. and  
Wiemann, S.  
EST (Poustka, A., Albert, R., Moosmayer, P., Schupp, I.,  
Wellenreuther, R., et al.)  
Unpublished (2003)  
**JOURNAL**  
Contact: MIPS  
MIPS Ingolstaedter Landstr. 1, D-85764 Neuberg, Germany  
This is the 5' sequence of the clone insert  
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer  
Research Center (DKFZ); Email: s.wiemann@dkfz-heidelberg.de;  
sequenced by DKFZ (German Cancer Research Center,  
Heidelberg/Germany) within the cDNA sequencing consortium of the  
German Genome Project.  
No 81 sequence available.  
This clone (DKFZP686H16201) is available at the RZPD in Berlin.  
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059  
Berlin-Charlottenburg, Germany; Email: clone@rzpd.de.  
**FEATURES**  
Location/Qualifiers  
1..735  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="DKFZP686H16201"  
/dev\_stage="adult"  
/lab\_host="DH10B"  
/clone\_lib="686 (synonym: hicc3)"  
/note="Vector: pT181EX2; Site\_1: SfiIA; Site\_2: SfiIB;  
cDNA-collection"  
**ORIGIN**  
Query Match 18.9%; Score 701.2; DB 13; Length 735;  
Best Local Similarity 99.3%; Pred. No. 2e-111; Indels 4; Gaps 1;  
Matches 714; Conservative 0; Mismatches 4;  
201 AGTGTATTCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 260  
2 AGTGTATTCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 61  
261 AGAATCAGTCT 320  
62 AGAATCAGTCT 121  
321 CATTTGACATTTCAAGATCATCAATCAATTTCAAGATTTGATGATGATGATGAT 380  
122 CATTTGACATTTCAAGATCATCAATCAATTTCAAGATTTGATGATGATGATGAT 181

Fri Mar 12 10:35:45 2004

us-10-084-817-30.rst

Page 13

Oy	381	GCTCAGACATCTGATGCTGCACTGTGCCAGACATCATCTGTCCTCTGAACTGAAATC	440
Db	182	GCTCAGACATCTGATGCTGCACTGTGCCAGACATCATCTGTCCTCTGAACTGAAATC	241
Oy	441	ACACCATGGAATGATTTTGAAGCTGCGAGAGAACTTGTGAAGGCGAAAGAGAGAGAAATGC	500
Db	242	ACACCATGGAATGATTTTGAAGCTGCGAGAGAACTTGTGAAGGCGAAAGAGAGAGAAATGC	301
Oy	501	GACTCGAAGCAGAAAGAAATGCGCTTACAGAGGATGACGATGATGAAGAGAGAGCAAGCC	560
Db	302	GACTCGAAGCAGAAAGAAATGCGCTTACAGAGGATGACGATGATGAAGAGAGAGCAAGCC	361
Oy	561	GCGAAGCGGCGCGCGAGCCCGACAGAAACGCGCTGCGGAGAAAGCAGAGAGAAAGAAATCCT	620
Db	362	GCGAAGCGGCGCGCGAGCCCGACAGAAACGCGCTGCGGAGAAAGCAGAGAGAAAGAAATCCT	421
Oy	621	TGGGACAGGTGACCGACCGACAGGTGGAGGTGAATGCCCAAGAAAGATGTGCTGACGAGAGAG	680
Db	422	TGGGACAGGTGACCGACCGACAGGTGGAGGTGAATGCCCAAGAAAGATGTGCTGACGAGAGAG	481
Oy	681	CCAGACAAACCAACCAAAACACTCAAGTGGAGAGGAGATGATGAGCCGCACTTCTTGAGC	740
Db	482	CCAGACAAACCAACCAAAACACTCAAGTGGAGAGGAGATGATGAGCCGCACTTCTTGAGC	541
Oy	741	GCCTGGCTGCGGTGAGAGAAAGCGCCCAAAAACGCTTCAAGAGAGCTCTTGAGACGCGACA	800
Db	542	GCCTGGCTGCGGTGAGAGAAAGCGCCCAAAAACGCTTCAAGAGAGCTCTTGAGACGCGACA	601
Oy	801	AGGATTTGAGACCCCAACATTAACAGATGCAAGTGTGCGCTCCCAAGCAGAGAAATGCATA	860
Db	602	AGGATTTGAGACCCCAACATTAACAGATGCAAGTGTGCGCTCCCAAGCAGAGAAATGCATA	661
Oy	861	ATGACACAGCAGAAAATGAAAATCAACCGAGAGAGAAAGAAAATGTAAGTGTCCCAAGAA	919
Db	662	ATGACACAGCAGAAAATGAAAATCAACCGAGAGAGAAATGTAAGTGTGTCCCAAGAA	719

Search completed: March 11, 2004, 14:23:47  
Job time : 6135.29 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 10, 2004, 08:38:17 ; Search time 5233.4 Seconds  
(without alignments)  
17011.249 Million cell updates/sec

Title: US-10-084-817-32  
Perfect score: 2054  
Sequence: 1 ccggccctgcctgtccgc.....cattgaaaaaaaaaaaaaa 2054

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues  
Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl.\*

- 1: gb\_ba.\*
- 2: gb\_btg.\*
- 3: gb\_in.\*
- 4: gb\_om.\*
- 5: gb\_ov.\*
- 6: gb\_pat.\*
- 7: gb\_ph.\*
- 8: gb\_pl.\*
- 9: gb\_pr.\*
- 10: gb\_ro.\*
- 11: gb\_sts.\*
- 12: gb\_sy.\*
- 13: gb\_un.\*
- 14: gb\_vi.\*
- 15: em\_ba.\*
- 16: em\_fun.\*
- 17: em\_hum.\*
- 18: em\_in.\*
- 19: em\_mu.\*
- 20: em\_om.\*
- 21: em\_or.\*
- 22: em\_ov.\*
- 23: em\_pat.\*
- 24: em\_ph.\*
- 25: em\_pl.\*
- 26: em\_ro.\*
- 27: em\_sts.\*
- 28: em\_un.\*
- 29: em\_vi.\*
- 30: em\_htg\_hum.\*
- 31: em\_htg\_inv.\*
- 32: em\_htg\_other.\*
- 33: em\_htg\_mus.\*
- 34: em\_htg\_pln.\*
- 35: em\_htg\_rod.\*
- 36: em\_htg\_mam.\*
- 37: em\_htg\_vrt.\*
- 38: em\_sy.\*
- 39: em\_htgo\_hum.\*
- 40: em\_htgo\_mus.\*
- 41: em\_htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2037.2	99.2	2043	9 BC031082	BC031082 Homo sapi
2	2034.4	99.0	2036	9 D87258	D87258 Homo sapien
3	2034.4	99.0	2075	6 E12965	E12965 cDNA encodi
4	2034	99.0	2036	6 AR095627	AR095627 Sequence
5	2025	98.6	2029	9 HSSERIPRT	Y07921 H.sapiens m
6	1766.6	86.0	1900	9 AK092476	AK092476 Homo sapi
7	1701.4	82.8	1894	6 AX747270	AX747270 Sequence
8	1701.4	82.8	1894	9 AK091944	AK091944 Homo sapi
9	1644	80.0	1705	6 AX015404	AX015404 Sequence
10	1644	80.0	1705	6 BD194550	BD194550 Human nuc
11	1455.2	70.8	1534	6 AX202090	AX202090 Sequence
12	1397	68.0	1423	9 AF070555	AF070555 Homo sapi
13	1330.2	64.8	2023	10 BC013516	BC013516 Mus muscu
14	1321	64.3	2005	10 AF179369	AF179369 Mus muscu
15	1321	64.3	2007	10 AK128916	AK128916 Mus muscu
16	1302.4	63.4	2010	10 AF172994	AF172994 Mus muscu
17	1301.8	63.4	2012	10 AF179370	AF179370 Rattus no
18	1162.6	56.6	1169	9 AF097709	AF097709 Homo sapi
19	953	46.4	1113	4 AF097707	AF097707 Bos tauru
20	717	34.9	79597	9 AF157623	AF157623 Homo sapi
21	717	34.9	111002	9 EX842242	EX842242 Homo sapi
22	717	34.9	183623	2 AL844554	AL844554 Mus muscu
23	717	34.9	191741	9 AL603764	AL603764 Human DNA
24	533.4	26.0	604	6 BD222460	BD222460 Drug-rela
25	493.4	24.0	495	6 AR273154	AR273154 Sequence
26	493.4	24.0	495	6 AR276735	AR276735 Sequence
27	493.4	24.0	495	6 AR407010	AR407010 Sequence
28	493.4	24.0	495	6 AX368187	AX368187 Sequence
29	479.4	23.3	539	6 I74320	I74320 Sequence 18
30	451.6	22.0	1797	6 AX338509	AX338509 Sequence
31	451.6	22.0	2543	9 AY280665	AY280665 Homo sapi
32	451.6	22.0	2554	9 AY040094	AY040094 Homo sapi
33	451.6	22.0	2585	9 BC034390	BC034390 Homo sapi
34	451.6	22.0	2586	9 BC035717	BC035717 Homo sapi
35	427.4	20.8	2446	10 AY156509	AY156509 Mus muscu
36	426.2	20.7	519	10 AF097706	AF097706 Cavia por
37	419	20.4	441	6 AR422696	AR422696 Sequence
38	419	20.4	441	6 BD118249	BD118249 Esf and e
39	414.6	20.2	1521	9 BC057765	BC057765 Homo sapi
40	414.6	20.2	1544	6 BD127576	BD127576 Primer fo
41	414.6	20.2	1544	9 AK075205	AK075205 Homo sapi
42	414.6	20.2	1553	6 BD063243	BD063243 Secreted
43	406.6	19.8	2040	6 BD260123	BD260123 Matrix-re
44	397.8	19.4	489	4 AF097708	AF097708 Oryctolag
45	391.4	19.1	1953	9 AY280666	AY280666 Homo sapi

ALIGNMENTS

RESULT 1  
BC031082 2043 bp mRNA linear PRI 08-OCT-2003  
LOCUS Homo sapiens protease, serine, 11 (IGF binding), mRNA (cDNA clone  
DEFINITION MGC:33808 IMAGE:5289018), complete cds.  
ACCESSION BC031082  
VERSION BC031082.1 GI:21410568  
KEYWORDS MGC.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 2043)  
Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,  
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,



Db 841 CGCTCTCAGAGCTGCGCGCGGAGAGTTCTGTGCGCATCGAAGCCCGTTTCCTT 900  
Qy 913 CAAAAACAGTCAACCGGATCGTAGACACCCACGAGCGGCGCAAGAGCTGGG 972  
Db 901 CAAAAACAGTCAACCGGATCGTAGACACCCACGAGCGGCGCAAGAGCTGGG 960  
Qy 973 CTCGCAACTCAGACATGGACTACATCAGACCGAGCGCATCATCAACTATGAAACTCG 1032  
Db 961 CTCGCAACTCAGACATGGACTACATCAGACCGAGCGCATCATCAACTATGAAACTCG 1020  
Qy 1033 GGAGGCGCGTAGTAACCTGGAGCGGTGAAGTATGGAATTAACACTTTGAAAGTGACA 1092  
Db 1021 GGAGGCGCGTAGTAACCTGGAGCGGTGAAGTATGGAATTAACACTTTGAAAGTGACA 1080  
Qy 1093 GCTGGAACTCTCTTGAATCCCATCTCATAGATTAAAGTTCTTCAACGAGTCCCAT 1152  
Db 1081 GCTGGAACTCTCTTGAATCCCATCTCATAGATTAAAGTTCTTCAACGAGTCCCAT 1140  
Qy 1153 GACGACAGGCGCAAGGAAAGCCATCAACCAAGAAAGATATTTGGTATCCGAATGATG 1212  
Db 1141 GACGACAGGCGCAAGGAAAGCCATCAACCAAGAAAGATATTTGGTATCCGAATGATG 1200  
Qy 1213 TCACCTACGCTCCAGCAAGCAAGAGCTGAAGACCGCGACTTCCAGACGCTG 1272  
Db 1201 TCACCTACGCTCCAGCAAGCAAGAGCTGAAGACCGCGACTTCCAGACGCTG 1260  
Qy 1273 ATCTCAGGAGCGTATATAATTGAAGTAATCTCTGATACCCAGCAGAGCTGGTGCTC 1332  
Db 1261 ATCTCAGGAGCGTATATAATTGAAGTAATCTCTGATACCCAGCAGAGCTGGTGCTC 1320  
Qy 1333 AAGGAAACGACGTCATAATCAGCATCAATGACAGCTCGTGCTCTCCGCCAATGATGC 1392  
Db 1321 AAGGAAACGACGTCATAATCAGCATCAATGACAGCTCGTGCTCTCCGCCAATGATGC 1380  
Qy 1393 AGGACGCTCAATTAAGGAAAGCAACCTGAACATGCTGGTCCGACAGGGGTATGAAGAT 1452  
Db 1381 AGGACGCTCAATTAAGGAAAGCAACCTGAACATGCTGGTCCGACAGGGGTATGAAGAT 1440  
Qy 1453 ATCATGATCAGAGTATCCGCAAGAAATGACCATAGCAGAGCATGAGCTGACATT 1512  
Db 1441 ATCATGATCAGAGTATCCGCAAGAAATGACCATAGCAGAGCATGAGCTGACATT 1500  
Qy 1513 CATGTTTCCCTCAAGACTCTCCGCTGATGACGATAGGACTCTGGGCTGCTGGAATA 1572  
Db 1501 CATGTTTCCCTCAAGACTCTCCGCTGATGACGATAGGACTCTGGGCTGCTGGAATA 1560  
Qy 1573 GGACACTCAGACTTTGACTGCCATTTGTTGTTTCACTGAGACTCCCTGSCCAACAG 1632  
Db 1561 GGACACTCAGACTTTGACTGCCATTTGTTGTTTCACTGAGACTCCCTGSCCAACAG 1620  
Qy 1633 AATCCTTCTTGATAGTTGCGAGGCAAAACAAATGTAATGTTGAGATCCGACGGCAGAAG 1692  
Db 1621 AATCCTTCTTGATAGTTGCGAGGCAAAACAAATGTAATGTTGAGATCCGACGGCAGAAG 1680  
Qy 1693 CTCTGCGCTTCTGATCTCTATGATGACGATGCTTTTCTTGCCAGCTTTGGGCCATTCT 1752  
Db 1681 CTCTGCGCTTCTGATCTCTATGATGACGATGCTTTTCTTGCCAGCTTTGGGCCATTCT 1740  
Qy 1753 TGCTTAGACAGTCAAGATTGCTCTCTCTCTTAACTGACTCATCAPCTTAGTCCAACTAA 1812  
Db 1741 TGCTTAGACAGTCAAGATTGCTCTCTCTCTTAACTGACTCATCTTAGTCCAACTAA 1800  
Qy 1813 TGCACTGATACAAATCGTAGTAGAAGAGCCCAAGGAGCCAGGATGGGACTGGTGC 1872  
Db 1801 TGCACTGATACAAATCGTAGTAGAAGAGCCCAAGGAGCCAGGATGGGACTGGTGC 1860  
Qy 1873 TGTGTTGCTTTTCTTCCAACTCAGACCCAAAGGTCATATGCAAGACCCCGGGTGGGT 1932  
Db 1861 TGTGTTGCTTTTCTTCCAACTCAGACCCAAAGGTCATATGCAAGACCCCGGGTGGGT 1920  
Qy 1933 GAGCGCTGGCTCTCAAGCGCGGAGTGGCTCTTTTAGGATCTCTTTGGAATGGGA 1992

1921 GAGCGCTGGCTTCTCAAAACGCGCGAAGTTGCTCTTTTAGGAATCTCTTTGGAATGGGA 1980  
Qy 1993 GCACGATGACTCTGAGTTGAGCTATTAAAGTACTCTTTACACATTGAAAAA 2052  
Db 1981 GCACGATGACTCTGAGTTGAGCTATTAAAGTACTCTTTACACATTGAAAAA 2040  
Qy 2053 AA 2054  
Db 2041 AA 2042

RESULT 2  
D87258  
LOCUS  
DEFINITION Homo sapiens mRNA for serin protease with IGF-binding motif,  
complete cds.  
ACCESSION D87258  
VERSION D87258.1 GI:1513058  
KEYWORDS serin protease with IGF-binding motif.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 2036)  
AUTHORS Ohno I., Hashimoto J., Takaoka K., Takahiro O., Okubo K. and Matsubara K.  
TITLE The cloning of a cDNA for novel serine protease with IGF-binding motif expressed in human osteoblast  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 2036)  
AUTHORS Ohno I.  
JOURNAL Direct Submission  
SUBMITTED (20-AUG-1996) Ikko Ohno, Institute for Molecular and Cellular Biology, Osaka University, Molecular Genetics; 1-3 Yamada-oka, Suita, Osaka 565, Japan  
[E-mail: Kousaku@imcb.osaka-u.ac.jp, Tel: 81-6-879-7992, Fax: 81-6-877-1922]

FEATURES  
source  
1. 2036  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="GS2422"  
/cell\_type="osteoblast"  
/tissue\_type="cancellous bone"  
/clone\_lib="lambda ZAP"  
/tissue\_lib="3 end-directed library"  
49..1491  
/note="GS2422 protein"  
/codon\_start=1  
/product="serin protease with IGF-binding motif"  
/protein\_id="BAA13322.1"  
/db\_xref="GI:1513059"

CDS  
1..1491  
/translation="MQIPRAALLPLLLLLAASPASQLSRAGSAPLAAGCPDRCEPA  
RCPPQPHCEGRARDACCEVCAGACGLOEGPCGEGLOCVVFPVFPASATVR  
RRAAGLCVCASSEPCSDANTYANLCQRAASRRSERLHREPVTVLQAGCQGOE  
DPNSLRKYNFIADVEKIAPAVHIELFKLPFPKREVPVAGSGFVSGDGLIVN  
AHVTNKHVAVELKNGATVEAKIKDVEKADIALIKIDHOGKLPVLLIGRSSELPG  
FVVAIGSPSELQNTVTGIVSTORGKEGLRSDMDYIOTDALINYGNSGGLPN  
LDGVIQNTLKTAGISFALPDSRIKKPLFTESHDRQAKGKAITKKYIGIRWMSLT  
SKAKELDRHDFPDVIGSYIIIEVDPPTAEAGGLKENDVIIISGQSVVANDVSD  
VIKRESLNVMVRRGNEDIMITVIEEDP"

ORIGIN  
Query Match 99.0%; Score 2034.4; DB 9; Length 2036;  
Best Local Similarity 100.0%; Pred. No. 4.8e-291;  
Matches 2035; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 CCGGCGCTTCCGCTGTCGCGCGCCACCGCGCGCGCGCAGAGTCGCCATGCAATCCG 60  
Db 1 CCGGCGCTTCCGCTGTCGCGCGCCACCGCGCGCGCGCAGAGTCGCCATGCAATCCG 60  
Qy 61 CCGGCGCTTCTTCTCCGCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 120



[illegible]

Db	1141	ACGGAGTCCCATGACCGACAGGCCAAGGAAAGCCATCACCAAGAGAGATATATTGGT	1200
Qy	1201	ATCCGAATGATGTCACTCACTCCAGCAAAAGCCAAAGAGCTGAAAGGACCGGCACCCGGAC	1260
Db	1201	ATCCGAATGATGTCACTCACTCCAGCAAAAGCCAAAGAGCTGAAGAGACCGGCACCCGGAC	1260
Qy	1261	TTCCACAGACGTGATCTCAGAGAGCGTATATAATTGAAGTAAATTCTGTATACCCACAGAGAA	1320
Db	1261	TTCCACAGACGTGATCTCAGAGAGCGTATATAATTGAAGTAAATTCTGTATACCCACAGAGAA	1320
Qy	1321	GCTGTGTGCTCAAGGAAACGACGTCAATTAATCAGCATCAATTGGACAGTCCGTGTCTCC	1380
Db	1321	GCTGTGTGCTCAAGGAAACGACGTCAATTAATCAGCATCAATTGGACAGTCCGTGTCTCC	1380
Qy	1381	GCCAAATGATGTACGACGAGTCAATTAAGAGGAAAGCACCCCTGAAACATGGTGTCTCCGACGG	1440
Db	1381	GCCAAATGATGTACGACGAGTCAATTAAGAGGAAAGCACCCCTGAAACATGGTGTCTCCGACGG	1440
Qy	1441	GGTAATGAAGATATCATATGATCAGATGATTTCCCGAAGAAATTGACCATAGCAGAGGCA	1500
Db	1441	GGTAATGAAGATATCATATGATCAGATGATTTCCCGAAGAAATTGACCATAGCAGAGGCA	1500
Qy	1501	TGAGCTGGACTTCATGTTTCCCTCAAGACTCTCCCGTGGATGACGGATGAGGACTCTGG	1560
Db	1501	TGAGCTGGACTTCATGTTTCCCTCAAGACTCTCCCGTGGATGACGGATGAGGACTCTGG	1560
Qy	1561	GCTGTGGAAATAGGACACTCAAGACTTTTGAATGCCATTTTGTGTTTCAGTGGAGACTC	1620
Db	1561	GCTGTGGAAATAGGACACTCAAGACTTTTGAATGCCATTTTGTGTTTCAGTGGAGACTC	1620
Qy	1621	CCTGGCCAAACAGAAATCCTTCTTGATAGTTTGACGCAAAAACAAATGTAAATGTTGCAGATC	1680
Db	1621	CCTGGCCAAACAGAAATCCTTCTTGATAGTTTGACGCAAAAACAAATGTAAATGTTGCAGATC	1680
Qy	1681	CGCAGCAGAAAGTCTGCCCTCTGTATCCTATGTATGACAGTGTGCTTTTTCTTTCGCCAGC	1740
Db	1681	CGCAGCAGAAAGTCTGCCCTCTGTATCCTATGTATGACAGTGTGCTTTTTCTTTCGCCAGC	1740
Qy	1741	TTGGGCCAATCTTTGCTTTAGACAGTCAGCATTTGTCTCCTCTTAACTGAGTCATCATCT	1800
Db	1741	TTGGGCCAATCTTTGCTTTAGACAGTCAGCATTTGTCTCCTCTTAACTGAGTCATCATCT	1800
Qy	1801	TAGTCCAACTAATGCAATGATCAATGCGGTAGATAGAAGAGCCCCACGGAGGCCAGGA	1860
Db	1801	TAGTCCAACTAATGCAATGATCAATGCGGTAGATAGAAGAGCCCCACGGAGGCCAGGA	1860
Qy	1861	TGCGACTGGTCTGTGTTTGTGCTTTTCTCCAAAGTCAGCACCCAAAGGTCAATGCACAGAGA	1920
Db	1861	TGCGACTGGTCTGTGTTTGTGCTTTTCTCCAAAGTCAGCACCCAAAGGTCAATGCACAGAGA	1920
Qy	1921	CCCCGGTGGTGAGCGCTGGCTTCTCAAACGGCCGGAAGTTGCCCTTTTATAGGAATCTCT	1980
Db	1921	CCCCGGTGGTGAGCGCTGGCTTCTCAAACGGCCGGAAGTTGCCCTTTTATAGGAATCTCT	1980
Qy	1981	TTGGAAATGGGACACGATCACTGTAGTTTGAGCTATTAAAGTACTTCTTACACA	2036
Db	1981	TTGGAAATGGGACACGATCACTGTAGTTTGAGCTATTAAAGTACTTCTTACAAA	2036

SECRET 3

57065  
E12965

SUBOT  
COSTA

## DEFINITION

11

**ACCESSION**

**VERSION**

## KEYWORDS

**SOURCE**

ORGANISM

**REFERENCE**

## REFERENCE ATTITUDE

## AUTHORS

E12965                      2075 bp    DNA    linear    PAT 27-APR-1998  
cDNA encoding novel physiologically active protein which have  
serine protease activity.

E12965.1 GI:3251794

JP 1997107980-A/1.

**Homo sapiens (human)**

**Homo sapiens**

**Eukaryota; Metazoa;**

Mammalia; Eutheria;

1 (bases 1 to 2075)

Hashimoto, A., Ooshiba







```
Db 485 CTTGACGCGGAGCTGCGGCCAAGGCGAGAGATCCCAACAGTTTGGCGCATAAATA 544
Qy 555 TAACTTTATCGCGGACGCTGGTGGAGAGATCGCCCTCGCGTGTTCATATCGAATTGTT 614
Db 545 TAACTTTATCGCGGACGCTGGTGGAGAGATCGCCCTCGCGTGTTCATATCGAATTGTT 604
Qy 615 TCGCAAGCTTCCGTTTTCTAAACGAGAGAGTCCCGTGGCTAGTGGTCTGGGTTTATTGT 674
Db 605 TCGCAAGCTTCCGTTTTCTAAACGAGAGAGTCCCGTGGCTAGTGGTCTGGGTTTATTGT 664
Qy 675 GTCGGAAGATGGAGCTGATCGTGACAAATGCCCCAGTGTGACCAACAAGCACCAGGTCAA 734
Db 665 GTCGGAAGATGGAGCTGATCGTGACAAATGCCCCAGTGTGACCAACAAGCACCAGGTCAA 724
Qy 735 AGTTGAGCTGAAGAACGGTGCCTTACGAAAGCCAAATCAAGAGATGGGATGAGAAAGC 794
Db 725 AGTTGAGCTGAAGAACGGTGCCTTACGAAAGCCAAATCAAGAGATGGGATGAGAAAGC 784
Qy 795 AGACATCGCACTCATCAAAATTTGACCAACAGGCAAGCTGCTGCTGCTGCTGGCGG 854
Db 785 AGACATCGCACTCATCAAAATTTGACCAACAGGCAAGCTGCTGCTGCTGCTGGCGG 844
Qy 855 CTCTCAGAGCTGCGGCGGGAGAGTTCGTGGTGCCTATCGGAAGCCGTTTTCCCTTCA 914
Db 845 CTCTCAGAGCTGCGGCGGGAGAGTTCGTGGTGCCTATCGGAAGCCGTTTTCCCTTCA 904
Qy 915 AAACACAGTCAACACCGGGATCGTGAGCACCACCCAGCGAGCGGCAAGAGCTGGGGCT 974
Db 905 AAACACAGTCAACACCGGGATCGTGAGCACCACCCAGCGAGCGGCAAGAGCTGGGGCT 964
Qy 975 CCGCACTCAGACATGACTATCACTCCAGCAGCGCCCATCACTCACTATGGAATCGGG 1034
Db 965 CCGCACTCAGACATGACTATCACTCCAGCAGCGCCCATCACTCACTATGGAATCGGG 1024
Qy 1035 AGCCCGCTTAGTAAACCTGAGCGTGAAGTGAATGGAATTAACACTTTGAAAGTGACAGC 1094
Db 1025 AGCCCGCTTAGTAAACCTGAGCGTGAAGTGAATGGAATTAACACTTTGAAAGTGACAGC 1084
Qy 1095 TGGATCTCTTTTCCATCCCATCTGTAAGATTAAGATTAAGATTAAGATTAAGATTAAG 1154
Db 1085 TGGATCTCTTTTCCATCCCATCTGTAAGATTAAGATTAAGATTAAGATTAAGATTAAG 1144
Qy 1155 CCGACAGCGCAAGGAAAGCCATCACAAGAAAGTATATGTTATCCGATGATGTC 1214
Db 1145 CCGACAGCGCAAGGAAAGCCATCACAAGAAAGTATATGTTATCCGATGATGTC 1204
Qy 1215 ACTCAGCTCCAGCAAGCAAGAGCTGAAGGACCGGACCGGACTTCCAGACGCTGAT 1274
Db 1205 ACTCAGCTCCAGCAAGCAAGAGCTGAAGGACCGGACCGGACTTCCAGACGCTGAT 1264
Qy 1275 CTAGGAGCGTATATATTAAGTAAATTCCTGATACCCAGCAGAGCTGGTGTCAA 1334
Db 1265 CTAGGAGCGTATATATTAAGTAAATTCCTGATACCCAGCAGAGCTGGTGTCAA 1324
Qy 1335 GGAAGAACGACGTCATAATCAGCATCAATGGACAGTCCGTTGCTCCGCCAATGATCTAG 1394
Db 1325 GGAAGAACGACGTCATAATCAGCATCAATGGACAGTCCGTTGCTCCGCCAATGATCTAG 1384
Qy 1395 CGAGCTCATTAAGGAAAGCAACCTGACATGTTGTTCCGACGGGTAAATGAAGATAT 1454
Db 1385 CGAGCTCATTAAGGAAAGCAACCTGACATGTTGTTCCGACGGGTAAATGAAGATAT 1444
Qy 1455 CATGATCAGTGAATCCCGAAGAAATGACCCATAGGACAGGCAAGTGGATTCATCA 1514
Db 1445 CATGATCAGTGAATCCCGAAGAAATGACCCATAGGACAGGCAAGTGGATTCATCA 1504
Qy 1515 TGTTCCTCAAGACTCTCCCGTGGATGACGATGAGGACTCTGGGCTGCTGGAATAGG 1574
Db 1505 TGTTCCTCAAGACTCTCCCGTGGATGACGATGAGGACTCTGGGCTGCTGGAATAGG 1564
Qy 1575 ACATCAAGACTTTGATCTGCCATTTTGTGTTGATGAGGAGCTCCCTGGCCCAACAGAA 1634
Db 1565 ACATCAAGACTTTGATCTGCCATTTTGTGTTGATGAGGAGCTCCCTGGCCCAACAGAA 1624
```

```
Qy 1635 TCCTTCTTGATAGTTTCAGGCAAAAACAAATGTAATGTTGCAGATCCGCGAGGAGAAGCT 1694
Db 1625 TCCTTCTTGATAGTTTCAGGCAAAAACAAATGTAATGTTGCAGATCCGCGAGGAGAAGCT 1684
Qy 1695 CTGCGCTTCTGATPCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1754
Db 1685 CTGCGCTTCTGATPCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1744
Qy 1755 CTTAGACAGTCAGCATTTGTCCTCTCTTAACTTAACTGAGTCATCATCTTAACTTAACT 1814
Db 1745 CTTAGACAGTCAGCATTTGTCCTCTCTTAACTTAACTGAGTCATCATCTTAACTTAACT 1804
Qy 1815 CAGTCGATCAATCGCTAGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1874
Db 1805 CAGTCGATCAATCGCTAGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1864
Qy 1875 TTTGTGCTTTTCTCAAGTCAGCACCACCAAGGTCAATGACAGAGACCCCGGTTGGTGA 1934
Db 1865 TTTGTGCTTTTCTCAAGTCAGCACCACCAAGGTCAATGACAGAGACCCCGGTTGGTGA 1924
Qy 1935 GCGCTGCTTCTCAACAGCGCGAAGTTGCTCTTTTAGGAATCTCTTTGGAATGGGAGC 1994
Db 1925 GCGCTGCTTCTCAACAGCGCGAAGTTGCTCTTTTAGGAATCTCTTTGGAATGGGAGC 1984
Qy 1995 ACGATGACTCTGAGTTGAGCTATTAAAGTACTTCTTACACATTTG 2039
Db 1985 ACGATGACTCTGAGTTGAGCTATTAAAGTACTTCTTACACATTTG 2029
```

```
RESULT 6
AK092476 1900 bp mRNA linear PRI 15-JUL-2002
LOCUS Homo sapiens cDNA FLJ35157 fis, clone PLACE011156, highly similar
DEFINITION AK092476
ACCESSION AK092476
VERSION AK092476.1 GI:21751082
KEYWORDS oligo capping; fis (full insert sequence)
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
```

```
REFERENCE 1 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS Kawakami,B., Sugiyama,A., Takemoto,M., Sugiyama,T., Irie,R.,
Oseuki,T., Sato,H., Wakamatsu,A., Ishii,S., Yamamoto,J., Isono,Y.,
Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K., Yamashita,H.,
Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H., Kanda,K.,
Wagatsuma,M., Murakawa,K., Kanehori,K., Takahashi-Fujii,A.,
Oshima,A., Suzuki,Y., Sugano,S., Nagahari,K., Masuho,Y., Nagai,K.
and Isogai,T.
```

```
TITLE NEDO human cDNA sequencing project
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1900)
AUTHORS Isogai,T. and Yamamoto,J.
TITLE Direct Submission
JOURNAL Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7
Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan
(E-mail:genomic@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
```

```
COMMENT NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology (RAB); cDNA library
construction: Helix Research Institute (HRI) (supported by Japan
Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
HRI, and Biotechnology Center, National Institute of Technology and
Evaluation; clone selection for full insert sequencing: HRI and
RAB; annotation: HRI and RAB.
```

```
FEATURES
source
1. 1900
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="PLACE011156"
/tissue_type="placenta"
```













Qy	1767	GCATTGTCTCTCCCTTTAACTAGTGCATCATCTTTAGTCCCACTAAATGCAGTCGATACAA	1820		
Db	1379	GCATTGTCTCTCCCTTTAACTAGTGCATCATCTTTAGTCCCACTAAATGCAGTCGATACAA	1438		
Qy	1827	TGCGTAGATAGAAAGAACGCCCAACGGGAGCCAGGATGGGACTGGTCTGTTTGTGCTTTTC	1886		
Db	1439	TGCGTAGATAGAAAGAACGCCCAACGGGAGCCAGGATGGGACTGGTCTGTTTGTGCTTTTC	1498		
Qy	1887	TCCAGTTCAGCACCCCAAGGTCATGACACAGAGACCCCGGGTGGGTGAGCGCTGGCTTCT	1946		
Db	1499	TCCAGTTCAGCACCCCAAGGTCATGACACAGAGACCCCGGGTGGGTGAGCGCTGGCTTCT	1558		
Qy	1947	CAAAAGCGCCGAAGTTGCCCTCTTTTAGGAATCTCTTTGGAAITGGGAGCAGATGACTCTG	2006		
Db	1559	CAAAAGCGCCGAAGTTGCCCTCTTTTAGGAATCTCTTTGGAAITGGGAGCAGATGACTCTG	1618		
Qy	2007	AGTTTGAGCTATTAAGTACTCTTTACACATTTGAAAAA	2054		
Db	1619	AGTTTGAGCTATTAAGTACTCTTTACACATTTGAAAAA	1666		
RESULT 10					
BD194550					
LOCUS	BD194550	1705 bp	DNA linear PAT 17-JUL-2003		
DEFINITION	Human nucleic acid sequence originating in normal ovarian tissue.				
ACCESSION	BD194550				
VERSION	BD194550.1	GI:33004292			
KEYWORDS	JP 2002510486-A/61.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	Speltz, T., Hintzman, B., Shcmitt, A., Pirarski, C., Duhl, E. and Rosenthal, A.				
TITLE	Human nucleic acid sequence originating in normal ovarian tissue				
JOURNAL	Patient: JP 2002510486-A 61 09-APR-2002;				
COMMENT	METAGEN GESELLSCHAFT FUER GENOME FORSCHUNG MBH OS Homo sapiens (human) PN JP 2002510486-A/61 PD 09-APR-2002 PF 31-MAR-1999 JP 2000542440 PR 03-APR-1998 DE 198 16 395.9 PI THOMAS SPEFT, BERND HINTZMAN, ARMIN SHCMITT, CHRISTIAN PIRARSKI, PI EDGAR DUHL, PI ANDRE ROSENTHAL PC C12N15/09, A61K38/00, A61K48/00, A61P35/00, C07K14/47, C07K16/30, PC C12N1/15, PC C12N1/19, C12N1/21, C12N5/10, C12Q1/68, G01N33/574, C12N15/00, PC A61K37/02, PC C12N5/00 CC Human nucleic acid sequence originating in normal ovarian CC				
FEATURES	FT	Key	Location/Qualifiers		
	FT	source	1..1705		
	FT		/organism='Homo sapiens (human)'		
			Location/Qualifiers		
ORIGIN			1..1705		
			/organism="Homo sapiens"		
			/mol_type="genomic DNA"		
			/db_xref="taxon:9606"		
Query Match 80.0%; Score 1644; DB 6; Length 1705;					
Best Local Similarity 99.9%; Pred. No. 2.3e-233;					
Matches 1666; Conservative 0; Mismatches 0; Indels 2; Gaps 2;					
Qy	387	CAGCAGGAGCCGGTGTGCGGACGACGCGCAACACCTACGCCAACCTGTGCCAGTGGC	446		
Db	1	CAGCAGGAGCCGGTGTGCGGACGAG-GAGCCCAACACCTACGCCAACCTGTGCCAGTGGC	59		
Qy	447	CGCCGACGACCGCGCTCCGAGAGAGCTTCACGCGCGCGGTCTATCGTCTGAGCGCGG	506		

```

QY 1587 TTTGACTGCCATTTTGTGTTTCAAGTGGAGACTCCCTGGCCAAACAGAAATCCTTCTTGATA 1646
Db 1199 TTTGACTGCCATTTTGTGTTTCAAGTGGAGACTCCCTGGCCAAACAGAAATCCTTCTTGATA 1258
QY 1647 GTTTGCAGGCAAAACAAATGTAATGTTGAGATCCCGCAGCAGCAAGCTCTGCCCTTCTGT 1706
Db 1259 GTTTGCAGGCAAAACAAATGTAATGTTGAGATCCCGCAGCAGCAAGCTCTGCCCTTCTGT 1318
QY 1707 ATCTATATGATGCAATGCTGCTTTTCTGCGAGCTTGGGCAATCTTGTAGACAGTCA 1766
Db 1319 ATCTATATGATGCAATGCTGCTTTTCTGCGAGCTTGGGCAATCTTGTAGACAGTCA 1378
QY 1767 GCATTGTCTCCTCTTAACTGAGTCATCATCTTAGTCCAACTAATGCAAGTCCGATACAA 1826
Db 1379 GCATTGTCTCCTCTTAACTGAGTCATCATCTTAGTCCAACTAATGCAAGTCCGATACAA 1438
QY 1827 TGGTATGATAGAAAGCCCAAGGAGGAGGAGATGGAGTGGTGTGTTGCTTTTC 1886
Db 1439 TGGTATGATAGAAAGCCCAAGGAGGAGGAGATGGAGTGGTGTGTTGCTTTTC 1498
QY 1887 TCCAAGTCAGACCCCAAGGTCATATGCAAGACACCCCGGTCGGTGAGCGCTGCTTCT 1946
Db 1499 TCCAAGTCAGACCCCAAGGTCATATGCAAGACACCCCGGTCGGTGAGCGCTGCTTCT 1558
QY 1947 CAACGGCCAGTGGCTCTTTTAGGAATCTCTTTGGAAATGGGAGCAGATGACTCTG 2006
Db 1559 CAACGGCCAGTGGCTCTTTTAGGAATCTCTTTGGAAATGGGAGCAGATGACTCTG 1618
QY 2007 AGTTTGAGCTATTAAGTACTTCTTACATTTGAAATGAAAAA 2054
Db 1619 AGTTTGAGCTATTAAGTACTTCTTACATTTGAAATGAAAAA 1666

RESULT 11
LOCUS AX202090 1534 bp DNA linear PAT 30-AUG-2001
DEFINITION Sequence 20 from Patent WO0153531.
ACCESSION AX202090
VERSION AX202090.1 GI:15391876
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Phippard,D., Vasanthakumari,G., Dotson,S. and Ma,X.J.
TITLE Osteoarthritis tissue derived nucleic acids, polypeptides, vectors,
and cells
JOURNAL Patent: WO 0153531-A 20 26-JUL-2001;
Pharmacia Corporation (US)
FEATURES
source
1..1534
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 70.8%; Score 1455.2; DB 6; Length 1534;
Best Local Similarity 99.3%; Pred. No. 1.8e-205;
Matches 1524; Conservative 0; Mismatches 3; Indels 7; Gaps 6;

QY 525 GGAAGATCCCAAGTTTGGCCATAAATATATCTTATCGGACGCTGGTGGAGAGAT 584
Db 1 GGAAGATCCCAAGTTTGGCCATAAATATATCTTATCGGACGCTGGTGGAGAGAT 60
QY 585 CGCCCTGCGGTGTTTCATATGAAATGTTTCGCAAGCTTCCGTTTCTAAACGAGAGGT 644
Db 61 CGCCCTGCGGTGTTTCATATGAAATGTTTCGCAAGCTTCCGTTTCTAAACGAGAGGT 120
QY 645 GCGGTGCTAGTGGTCTGGTCTTATGTTGCGAAGATGGATGATCGTGAATAATGC 704
Db 121 GCGGTGCTAGTGGTCTGGTCTTATGTTGCGAAGATGGATGATCGTGAATAATGC 180

```

```

QY 705 CCACTGTGTGACCAACAAGACACCGGCTCAAAGTTGAGTGAAGAACGGTGCCTTACGA 764
Db 181 CCACTGTGTGACCAACAAGACACCGGCTCAAAGTTGAGTGAAGAACGGTGCCTTACGA 240
QY 765 AGCCAAAATCAAGGATGTGGATGAGAAAGCAGACATCGCACTCATCAAAAATGACCA 824
Db 241 AGCCAAAATCAAGGATGTGGATGAGAAAGCAGACATCGCACTCATCAAAAATGACCA 300
QY 825 GGGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 884
Db 301 GGGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
QY 885 GGTCCGCAATCGAAGCCCGTTTCCCTTCAAAAACACAGTCACCAACCGGGATCGTGAGCAC 944
Db 361 GGTCCGCAATCGAAGCCCGTTTCCCTTCAAAAACACAGTCACCAACCGGGATCGTGAGCAC 420
QY 945 CACCCAGGAGCGGCAAGAGTGGGGTCCGCAACTCAGACATGAGACTACATCCAGAC 1004
Db 421 CACCCAGGAGCGGCAAGAGTGGGGTCCGCAACTCAGACATGAGACTACATCCAGAC 480
QY 1005 CGACGCCATCATCAACTATGGGAACTCCGGGAGGCGCTTAGTAAACCTTGACGGTGAA 1062
Db 481 CGACGCCATCATCAACTATGGGAACTCCGGGAGGCGCTTAGTAAACCTTGACGGTGAA 540
QY 1063 GTGATTGGAATTAACACTTTGAAAGTGACAGCTGGAAATCTCCTTTGCAATCCCATCTGAT 1122
Db 541 GTGATTGGAATTAACACTTTGAAAGTGACAGCTGGAAATCTCCTTTGCAATCCCATCTGAT 600
QY 1123 AAGATTAAAAAGTTTCTCAGCGAGTCCCATGACCGACAGGCCCAAGGAAAGCCATCAC 1182
Db 601 AAGATTAAAAAGTTTCTCAGCGAGTCCCATGACCGACAGGCCCAAGGAAAGCCATCAC 660
QY 1183 AAGAAAGATATATTGGTATCCGAATGATGTCACTCAGCTCCAGCAAGCCAAAGAGCTG 1242
Db 661 AAGAAAGATATATTGGTATCCGAATGATGTCACTCAGCTCCAGCAAGCCAAAGAGCTG 720
QY 1243 AAGGACCGGCACCGGACCTTCCAGACGCTGATCTCAGGAGCGTATATAATTAAGTAAT 1302
Db 721 AAGGACCGGCACCGGACCTTCCAGACGCTGATCTCAGGAGCGTATATAATTAAGTAAT 780
QY 1303 CTTGATACCCAGCAGAGAGCTGTG-TGCTCAAGGAAAGAGAGCTCATATCAGATCAA 1361
Db 781 CTTGATACCCAGCAGAGAGCTGTG-TGCTCAAGGAAAGAGAGCTCATATCAGATCAA 840
QY 1362 TGGACAGTCCGTGCTCCGCCAATGATGTCAAGCAGCTG-TCAATTAAGGAGGAAAGCACCC 1420
Db 841 TGGACAGTCCGTGCTCCGCCAATGATGTCAAGCAGCTG-TCAATTAAGGAGGAAAGCACCC 900
QY 1421 TGAACATGTTGTTCCGAGGCTTAATGAAGATATCATGATCAGATGATCCGAGGAA 1480
Db 901 TGAACATGTTGTTCCGAGGCTTAATGAAGATATCATGATCAGATGATCCGAGGAA 960
QY 1481 TTGACCCATAGGAGGATGAGCTGGAGCTTCAATGTTTCCCTCAAAGAGCTCTCCCGT-G 1539
Db 961 TTGACCCATAGGAGGATGAGCTGGAGCTTCAATGTTTCCCTCAAAGAGCTCTCCCGTGG 1020
QY 1540 GATGACGATAGGAGCTTGGGCTGCTGGAATAGGACATCAAGACTTTTACCTGCCATT 1599
Db 1021 GATGACGATAGGAGCTTGGGCTGCTGGAATAGGACATCAAGACTTTTACCTGCCATT 1080
QY 1600 TTGTTTGTTCAGTGGAGACTCCCTGGCCCAAGAGATCTTCTTATAGTTTTCAGGCAAA 1659
Db 1081 TTGTTTGTTCAGTGGAGACTCCCTGGCCCAAGAGATCTTCTTATAGTTTTCAGGCAAA 1140
QY 1660 ACAATGTAATGTTGAGATCCGAGGAGAGCTCTG-CCCTTCTGATCTGATGATG 1718
Db 1141 ACAATGTAATGTTGAGATCCGAGGAGAGCTCTG-CCCTTCTGATCTGATGATG 1200
QY 1719 CAGTGTCTTTTCTTGGCAGCTTGGGCAATCTTGTCTAGACAGTCAAGCTTTTCTCTCC 1778
Db 1201 CAGTGTCTTTTCTTGGCAGCTTGGGCAATCTTGTCTAGACAGTCAAGCTTTTCTCTCC 1260

```



```
QY 1779 TCCTTTAACTGAGTCATCATCTTAGTCCAACTAATGAGTGCAGTCGATACAATG-CGTAGATAG 1837
Db 1261 TCCTTTAACTGAGTCATCATCTTAGTCCAACTAATGAGTGCAGTCGATACAATGCGGTAGATAG 1320
QY 1838 AAGAAGCCCAACGGGAGCCAGGATGGAGCTGGTCTGCTGTTTGTGCTTTTCTCCAACTCAGC 1897
Db 1321 AAGAAGCCCAACGGGAGCCAGGATGGAGCTGGTCTGCTGTTTGTGCTTTTCTCCAACTCAGC 1380
QY 1898 ACCCAAAGTCAATGACACAGACCCCGGTTGGGTGAGCGTGGCTTCTCAAACGCCGA 1957
Db 1381 ACCCAAAGTCAATGACACAGACCCCGGTTGGGTGAGCGTGGCTTCTCAAACGCCGA 1440
QY 1958 AGTTGCTCTTTTAGGAATCTCTTTGGAATGGAGCAGATGACTCTGAGTTGAGCTA 2017
Db 1441 AGTTGCTCTTTTAGGAATCTCTTTGGAATGGAGCAGATGACTCTGAGTTGAGCTA 1500
QY 2018 TTAAGTACTCTTACACATTGAAAAA 2051
Db 1501 TTAAGTACTCTTACACATTGAAAAA 1534

RESULT 12
AF070555
LOCUS AF070555 1423 bp mRNA linear PRI 05-AUG-1998
DEFINITION Homo sapiens clone 24795 mRNA sequence.
ACCESSION AF070555
VERSION AF070555.1 GI:3387920
KEYWORDS FLI CDNA.
SOURCE Homo sapiens
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Carnivora; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1423)
AUTHORS Anderson,B., Wentland,M.A., Ricafrente,J.Y., Liu,W. and Gibbs,R.A.
TITLE A 'double adaptor' method for improved shotgun library construction
JOURNAL Anal. Biochem. 236 (1), 107-113 (1996)
MEDLINE 96207227
PUBMED 8619474
REFERENCE 2 (bases 1 to 1423)
AUTHORS Yu,W., Anderson,B., Worley,K.C., Muzny,D.M., Ding,Y., Liu,W.,
Ricafrente,J.Y., Wentland,M.A., Lennon,G. and Gibbs,R.A.
TITLE Large-scale concatenation cDNA sequencing
JOURNAL Genome Res. 7 (4), 353-358 (1997)
MEDLINE 9110174
PUBMED 9110174
REFERENCE 3 (bases 1 to 1423)
AUTHORS Yu,W. and Gibbs,R.A.
TITLE Direct Submission
JOURNAL Submitted (05-JUN-1998) Molecular and Human Genetics, Baylor
College of Medicine, One Baylor Plaza S930, Houston, TX 77030, USA
FEATURES
source
Location/Qualifiers
1. 1423
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="I.M.A.G.E. Consortium clone ID 24795"
/sex="female"
/tissue_type="brain"
/clone_lib="INIB"
/dev_stage="infant"
/notes="similar to Homo sapiens mRNA for serine protease
with GenBank Accession Number Y07921"

ORIGIN
Query Match 68.0%; Score 1397; DB 9; Length 1423;
Best Local Similarity 99.9%; Pred. No. 7.2e-197;
Matches 1408; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 646 CCGGTGGCTAGTGGGTCTGGGTTTATTGTGCGGAAGATGGACTGATCGTGACAAATGCC 705
Db 1 CCGGTGGCTAGTGGGTCTGGGTTTATTGTGCGGAAGATGGACTGATCGTGACAAATGCC 60
QY 706 CAGTGTGTACCAACAAAGACCCGGGTCAAAGTTGAGTGTGAAGACGGTGCCTACGAA 765
```

```
Db 61 CAGTGTGTACCAACAAAGACCCGGGTCAAAGTTGAGTGTGAAGACGGTGCCTTACGAA 120
QY 766 GCCAAATCAAGGATGTGGATGAAAGCAGACATCGCACTCATCAAATTTGACCCAG 825
Db 121 GCCAAATCAAGGATGTGGATGAAAGCAGACATCGCACTCATCAAATTTGACCCAG 180
QY 826 GGCAAGTGTCTGTCTGCTGCTTGGCGCTCTCAGAGCTCGCGCGGAGAGTTCGTG 885
Db 181 GGCAAGTGTCTGTCTGCTGCTTGGCGCTCTCAGAGCTCGCGCGGAGAGTTCGTG 240
QY 886 GTGCGCATCGGAAGCCCGTTTCCCTTCAAACACAGTACACACCGGATCGTGAGACC 945
Db 241 GTGCGCATCGGAAGCCCGTTTCCCTTCAAACACAGTACACACCGGATCGTGAGACC 300
QY 946 ACCCAGCGGCGGCAAGAGCTGGGGCTCCGCAACTCAGACATGAGCTACATCCAGACC 1005
Db 301 ACCCAGCGGCGGCAAGAGCT-GGGCTCCGCAACTCAGACATGAGCTACATCCAGACC 359
QY 1006 GAGCCATCATCAACTATGGAACCTCGGAGGCCCGTTAGTAAACCTGACGCTGAAGTG 1085
Db 360 GAGCCATCATCAACTATGGAACCTCGGAGGCCCGTTAGTAAACCTGACGCTGAAGTG 419
QY 1066 ATTGGAATTAACACTTTGAAAGTGACAGCTGGAATCTCTTTTGCAATCCCATCTGATAAG 1125
Db 420 ATTGGAATTAACACTTTGAAAGTGACAGCTGGAATCTCTTTTGCAATCCCATCTGATAAG 479
QY 1126 ATTAAGATTTCTCAGCGAGTCCCATGACGACAGGCGCAAGAAAGAAAGCCATCACCAAG 1185
Db 480 ATTAAGATTTCTCAGCGAGTCCCATGACGACAGGCGCAAGAAAGAAAGCCATCACCAAG 539
QY 1186 AAGAAGTATATTGGTATCCGAATGATGTCACCTCAGCTCCAGCAAGCCAAAGAGCTGAAG 1245
Db 540 AAGAGTATATTGGTATCCGAATGATGTCACCTCAGCTCCAGCAAGCCAAAGAGCTGAAG 599
QY 1246 GACCGGCAACCGGAGCTTCCAGACGCTGATCTCAGAGCGGTATATAATTGAAGTAATTCCT 1305
Db 600 GACCGGCAACCGGAGCTTCCAGACGCTGATCTCAGAGCGGTATATAATTGAAGTAATTCCT 659
QY 1306 GATACCCCGCAGAGAGCTGGTGGTCTCAAGGAAACAGCGTCTATATCAGCATCAATGGA 1365
Db 660 GATACCCCGCAGAGAGCTGGTGGTCTCAAGGAAACAGCGTCTATATCAGCATCAATGGA 719
QY 1366 CAGTCCGTGCTCCGCAATGATGTACGACGCTCAATTAAGGAAAGACCCCTGAAC 1425
Db 720 CAGTCCGTGCTCCGCAATGATGTACGACGCTCAATTAAGGAAAGACCCCTGAAC 779
QY 1426 ATGTGCTCCGCGAGGGTAAATGAAGATATCATGATCAGATGATCCCGAAGAAATGAC 1485
Db 780 ATGTGCTCCGCGAGGGTAAATGAAGATATCATGATCAGATGATCCCGAAGAAATGAC 839
QY 1486 CCATAGCAGAGGATGAGCTGGACTTCATGTTTCCCTCAAAGACTCTCCCGTGGATGAC 1545
Db 840 CCATAGCAGAGGATGAGCTGGACTTCATGTTTCCCTCAAAGACTCTCCCGTGGATGAC 899
QY 1546 GGATGAGGACTCTGGCTGCTGGAAATAGACACTCAAGACTTTTGTACTGCAATTTCTTT 1605
Db 900 GGATGAGGACTCTGGCTGCTGGAAATAGACACTCAAGACTTTTGTACTGCAATTTCTTT 959
QY 1606 GTTCAGTGGAGACTCCCTGGCCAAACAGAAATCTTCTTGTAGTATTTCAGGCAAAACAAAT 1665
Db 960 GTTCAGTGGAGACTCCCTGGCCAAACAGAAATCTTCTTGTAGTATTTCAGGCAAAACAAAT 1019
QY 1666 GTAAATTTGCAGATCCCGCAGGACAGAGCTCTGCCCTCTGTATCTCTATGTATGATGAGTGTG 1725
Db 1020 GTAAATTTGCAGATCCCGCAGGACAGAGCTCTGCCCTCTGTATCTCTATGTATGATGAGTGTG 1079
QY 1726 CTTTTCCTTCCAGCTTGGGCAATCTTTGCTTAGACAGTACAGATTTGCTCTCTCTTTA 1785
Db 1080 CTTTTCCTTCCAGCTTGGGCAATCTTTGCTTAGACAGTACAGATTTGCTCTCTCTTTA 1139
QY 1786 ACTGAGTCATCATCTTAGTCCAACTAATGCGAGTCGATCAATGCGGTAGATAGAGAGCC 1845
```



R. M. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAK Plate: 25 Row: C Column: 8

This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 9625026.

Location/Qualifiers

1. .2023

/organism="Mus musculus"

/mol\_type="mRNA"

/strain="FVB/N"

/db\_xref="taxon:10090"

/clone="MGC:13188 IMAGE:4235912"

/tissue\_type="kidney, normal, 5 month old male mouse."

/clone\_lib="NCI-CCAP\_Kid14"

/lab\_host="DH10B"

/note="Vector: pCMV-SPORT6"

1. .2023

/gene="Prss11"

/note="synonym: HTRA"

/db\_xref="LocusID:56213"

/db\_xref="MGI:1929076"

40\_1482

/codon\_start=1

/product="protease, serine, 11 (Igf binding)"

/protein\_id="AAH3516.1"

/db\_xref="GI:15488756"

/db\_xref="LocusID:56213"

/translation="MQSLRTVLLSLLLLLAAPSLPGTGRSPAATVCPPEHCDPT RCAPPTDCGEVRVDRACGCGCEGALGAGLQEGPCGEGLCQVFPVGPASATVR DPAAGSCVASSPFCVGSQDTYTNQLRAASRRSKLRQPPVILQRCAGCGQCE RNSLRHKYFNPAWVEKIAAPVHIEURLKLPFSKREVPVASGQFVISEDGLIVN AHVYTKNRKSVLEIKNGATYEAKIKVDDEKADIALIKDHQGLKLVLLGRSSELRAP FVVAIGKPSVLTQVTVSTQRCGEKLEGLRNSDMWYICTDAIINYGNSGGLPVN LDGEVIGINILKVTAGISFAPSKIKKFLQSHDRQAKGKAVTKKVIKIGIRNMSLTS SKALKEDRRHDFEDVSGSIVIEVPTPRAEGKLENDVLIISNGQSVVTADVSD VIKKEKTLNMRHDFEDVSGSIVIEVPTPRAEGKLENDVLIISNGQSVVTADVSD 148..375

/note="1B; Region: Insulin growth factor-binding protein homologues"

/db\_xref="CDD:smart00121"

499..1479

/note="Dego; Region: Trypsin-like serine proteases, typically periplasmic, contain C-terminal PDZ domain [Pozstranslational modification, protein turnover, chaperones]"

/db\_xref="CDD:COG0265"

h 64.8; Score 1330.2; DB 10; Length 2023;

Similarity 81.9%; Pred.No.5.1e-187;

662; Conservative 0; Mismatches 338; Indels 29; Gaps 10;

9 CCGCGCGCGCAGAGTCGCCATGCAGATCCCGCGCGCGCTCTTCTCCGCGTCTCTCTGC 88

0 CCGTGGCGTGGAGTGGTATGATGCTCCGTCGTCGTCACCGTCTCTTGTCTACTGC 79

9 TGCTGTGGCGCGCGCGCTCGGGCGAGCTGTCCCGCGCGCGCGCTCGGCGCTTGG 148

0 TGCTGTAGCGGCTCTTCTTGGGGTGTGGCTCGGGGACCGCGCGCTCGGCGGAGCTG 139

9 CCGCGGGTGGCCAGACCGCTCGAGACCGCGCGCTGCGCGCGCGAGCGAGCACTGG 208

0 CCACGTCTGTCCGAGCACTGGATCCACCGCTGCGCGCGCGCGCACTGG 199

9 AGGGCGCGCGCGCGCGGACCGTCCGGCTGTCTGAGGTGTGGCGCGCGCGCGGCG 268

0 AGGTGGCGCGCTCGCGACCGCTGCGCTGTGAGGTGTGGCGCGCGCTCGAGGCTG 259

9 CCGGTGGCGCTCGAGAGGCGCGTGGCGCGAGGCGCTGAGTGTGGTGGTCCCTTCG 328

0 CAGGTGGCGCTCGAGAGGCTCGTGGCGAGGCGCTGCAATCGTAGTGGTCCCTTCG 319

Qy	1409	GGGAAAGCACCTCTGAACATGGTGGTCCGACAGGGGTATGAAGATATCATGTACAGTGA	1468
Db	1400	AGGAGAACACCTCGTGAACATGGTGTCCGACAGGGCAATGAAGACATTTGTATCCGTGA	1459
Qy	1469	TTCCGGAAGAAATTGACCCCATAGCCAGAGGCGATGAGCTGGACTTCATGTTCCCTCAAAG	1528
Db	1460	TTCTCTGAAGAAATCGACCCCTAGCCAGAGGCGAGGACGACTTCATGTTCCCTCAAAG	1519
Qy	1529	ACTCTCCCGTGGATGACGGATGAGGACTCTGGGCTGTGGNATAGGACACTCAAGACTTT	1588
Db	1520	A--CTCCAGGGGATGGCGCATGAGAA--CTGAGCTGGTGACGCGAGGACCCGGGACCTT	1575
Qy	1589	TGACTGCCATTTGTTTGTTCAGTGGAGACTCCCTTGCCCAACAGAATCCTTCTTGATAGT	1648
Db	1576	TGTCCAAATTTGCTGTTTCAGGGAAACACCCCT--GCCGACAGAACTCTTCTTGATAGT	1633
Qy	1649	TTGCAGGCAAAACAAATGTAATGTTTCAGATCCGACGCGAGAAGCTCGCCCTCTGTAT	1708
Db	1634	TTGCGGGCAAAACAAATGTAATGTTGCGAGGTGCACAGGCGAGNCTCTG--CTTCTGTAT	1692
Qy	1709	CCTATGTATGACAGTGTGCTTTTTCTTTGCCAGCTTGGGCCATCTTGCTTAGACAGTCAGC	1768
Db	1693	GCTATGTATGACGCGTCTTTTCTTTACAAGCTTGGGCTGTCTCGTCTTACACAGTCAAC	1752
Qy	1769	ATTGTCTCTCTTTAACTGAGTCATCATCTTAGTCCAACTAATGCAGTGATACAAATG	1828
Db	1753	ATTGTCTCTTCCCTTAGC-----CTAAGTTGCTGACTAATGCAGTTGACGGAATGC	1804
Qy	1829	CGTAGATAGAGAAGCCCCACGCGGACCGAGGATGGGACTGGTCTGTGTTTGCTCTTTT--CT	1887
Db	1805	GTAGGCAAGAGGAAGGTCCTC---GCAGCCATGGTCTTGGCGGTGTTTACACTTTTCT	1860
Qy	1888	CCAACTCAGCACCCCAAAAG---TCAATGCACAGACCCCGGGTGGGTGAGCGCTGGCTT	1944
Db	1861	CCGAGTCGCACTCAGAGGAAATTGATGCCCCGAGACCAAGGTTGGGTGA--CCCTGGCTT	1919
Qy	1945	CTCAACCGCCGGAAGTTGGCTCTTTTAGGATCTCTTTGGAAATTTGGGACGAGTACATC	2004
Db	1920	CTGAGATGCCCAAAATGCTCTT-----AGGAAATCTCTGAAACCCGGAGTACAGTGACTT	1974
Qy	2005	TGAGTTTGAGCTATTAAAGTACTCTTTACACATTTGAAAAAATAAAAAA	2053
Db	1975	TGAGTTTCAGCTATTAAATACTTTCTTACGTGAAAAAATAAAAAA	2023

RESULT 14  
 AF179369  
 LOCUS  
 DEFINITION  
 AF179369  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 FEATURES

2005 bp linear  
 Mus musculus insulin-like growth factor binding protein 5 protease  
 mRNA, complete cds.  
 AF179369.1 GI:5815458  
 Mus musculus (house mouse)  
 Mus musculus  
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 2005)  
 Hourvitz,A., Hennebold,J.D., King,G., Negishi,H., Erickson,G.F.,  
 Roby,J.A., Mayo,K.E. and Adashi,E.Y.  
 Mouse insulin-like growth factor binding protein 5-directed  
 endopeptidase: structural assessment, evolutionary analysis,  
 ovarian expression, hormonal regulation and cellular localization  
 Unpublished  
 2 (bases 1 to 2005)  
 Hourvitz,A., Hennebold,J.D. and Adashi,E.Y.  
 Direct Submission  
 Submitted (20-AUG-1999) Reproductive Sciences, University of Utah,  
 546 Chipera Way, Salt Lake City, UT 84108, USA  
 Location/Qualifiers  
 1..2005  
 /organism="Mus musculus"  
 /mol\_type="mRNA"







GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

M nucleic - nucleic search, using sw model

Run on: March 10, 2004, 08:01:57 ; Search time 499.903 Seconds  
(without alignments)  
17454.974 Million cell updates/sec

Title: US-10-084-817-32

Perfect score: 2054

Sequence: 1 cgggctccgtccgtccgc.....cattgaaaaaaaaaaaaa 2054

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: N\_Geneseq\_29Jan04.\*
- 2: Geneseq1980s.\*
- 3: Geneseq1990s.\*
- 4: Geneseq2000s.\*
- 5: Geneseq2001as.\*
- 6: Geneseq2001bs.\*
- 7: Geneseq2002as.\*
- 8: Geneseq2003as.\*
- 9: Geneseq2003bs.\*
- 10: Geneseq2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2037.4	99.2	2157	6	ABQ54255 Human ova
2	2034.4	99.0	2036	6	ABV78051 Hypoxia-x
3	2034.4	99.0	2075	2	AAT75444 cDNA enco
4	2034.4	99.0	2036	2	AAT75444 Homo sapi
5	2010.8	97.9	2205	2	AAV29540 Human pre
6	1780	86.7	1855	3	AAF21747 Human bre
7	1701.4	82.8	1894	9	ADB62641 Human cDN
8	1644	80.0	1705	2	AZ41293 Human nor
9	1455.2	70.8	1534	4	AH23090 Osteoarth
10	1440	70.1	1440	6	AD44188 Human TAN
11	1191.8	58.0	1443	3	AZ229176 Recombina
12	1043.8	50.8	1185	3	AZ229195 cDNA enco
13	759.8	37.0	1389	7	ABX77560 Different
14	635	30.9	669	6	ABQ59513 Human col
15	558.4	27.2	900	4	AAL24929 Human bre
16	530	25.8	895	4	AAL24918 Human bre
17	493.4	24.0	495	6	ABK38859 cDNA enco
18	493.4	24.0	495	7	ACAL1188 Human lun
19	493.4	24.0	495	7	ACA02374 Lung can
20	478.2	23.3	487	3	AAA30449 Human bra
21	451.6	22.0	1797	6	AAD23855 Human pro
22	451.6	22.0	2094	6	ABK51497 cDNA enco
23	451.6	22.0	2543	7	ACF12810 Human PRS

24	451.6	22.0	2577	3	AAA57359
25	447.2	21.8	607	2	AXX29130
26	447	21.8	1583	7	ACF16993
27	446	21.7	490	8	ACH26037
28	427.4	20.8	2450	7	ACF12809
29	414.6	20.2	1544	4	AAK94547
30	414.6	20.2	1553	2	AAV43618
31	406.6	19.8	2040	3	AZ52362
32	406.6	19.8	2040	6	AAI42458
33	402.2	19.6	457	8	ACH13892
34	395.8	19.3	1044	6	AD44189
35	395.8	19.3	1563	3	AAA57361
36	391.4	19.1	1953	7	ACF12811
37	389	18.9	705	4	AAI24879
38	386.4	18.8	1695	4	AAI4885
39	385.6	18.8	1788	4	AAI26848
40	384	18.7	384	7	ACF12817
41	381.8	18.6	1859	4	AD10126
42	381.8	18.6	1859	6	ABL35059
43	365.6	17.8	1897	7	ACF12815
44	352.8	17.2	412	7	ABX39463
45	331.8	16.2	1193	6	ABQ30909

## ALIGNMENTS

### RESULT 1

ABQ54255

ID ABQ54255 standard; cDNA, 2157 BP.

XX AC ABQ54255;

XX XX

DT 22-AUG-2002 (first entry)

XX XX

DE Human ovarian antigen HKADF15 cDNA, SEQ ID NO:135.

XX XX

Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;  
ovarian cancer; breast cancer; tumour; reproductive system disorder;  
infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;  
PCOS; ovarian cyst; dysmenorrhea; endocrine disorder; infection;  
inflammatory condition; immune disorder; blood disorder;  
cardiovascular disorder; respiratory disorder; neurological disorder;  
gastrointestinal disorder; urinary system disorder; drug screening;  
gene therapy; chromosome mapping; forensic analysis;  
antibody preparation; cytostatic; immunomodulatory; neuroprotective;  
antiinflammatory; gynaecological; reproductive; chromosome 10q25.3-26.2;  
gene; ss.

XX OS Homo sapiens.

XX PN WO200200677-A1.

XX XX

PD 03-JAN-2002.

XX XX

PF 07-JUN-2001; 2001WO-US018569.

XX XX

PR 07-JUN-2000; 2000US-0209467P.

XX XX

PA (HUMA-) HUMAN GENOME SCI INC.

XX XX

PI Birse CE, Rosen CA;

XX XX

DR WP; 2002-147878/19.

XX XX

DR P-PSDB; ABP41178.

XX XX

Isolated nucleic acid molecules encoding novel ovarian polypeptides,  
useful in the prevention, treatment and diagnosis of cancer (e.g. ovarian  
cancer), immune disorders, cardiovascular disorders and neurological  
diseases.

Claim 1; SEQ ID NO 135; 2922pp; English.







QY 721 AAGCAGCGGTCAAGTTGAGCTGAGAAACGGTGCCTACTTACGAGCCAAATCAAGAT 780  
 Db 721 AAGCAGCGGTCAAGTTGAGCTGAGAAACGGTGCCTACTTACGAGCCAAATCAAGAT 780  
 QY 781 GTGGATGAGAAAGCAGACATCGCACTCATCAAAATTGACCAACAGGGCAAGCTGCTGTC 840  
 Db 781 GTGGATGAGAAAGCAGACATCGCACTCATCAAAATTGACCAACAGGGCAAGCTGCTGTC 840  
 QY 841 CTGCTGTTGGCCGCTCCTCAGAGCTGGCGGGGAGAGATTGCTGGTGGCCATCGGAAGC 900  
 Db 841 CTGCTGTTGGCCGCTCCTCAGAGCTGGCGGGGAGAGATTGCTGGTGGCCATCGGAAGC 900  
 QY 901 CCGTTTTTCCCTTCAAAACACAGTCAACACCGGGATCGTGAGCAACCCAGCGAGCGGC 960  
 Db 901 CCGTTTTTCCCTTCAAAACACAGTCAACACCGGGATCGTGAGCAACCCAGCGAGCGGC 960  
 QY 961 AAGAGCTGGGGCTCCGCAATCAGACATGAGTACATCCAGACGAGCCCATCATCAAC 1020  
 Db 961 AAGAGCTGGGGCTCCGCAATCAGACATGAGTACATCCAGACGAGCCCATCATCAAC 1020  
 QY 1021 TATGGAACTCGGGAGGCCCTTATGTAACCTGGACGGTGAAGTGATTGGAATTAACACT 1080  
 Db 1021 TATGGAACTCGGGAGGCCCTTATGTAACCTGGACGGTGAAGTGATTGGAATTAACACT 1080  
 QY 1081 TTGAAGTGACAGTGGAAATCTCTTTCATCCATCCATCTGATGAATTAAGTTCCCTC 1140  
 Db 1081 TTGAAGTGACAGTGGAAATCTCTTTCATCCATCCATCTGATGAATTAAGTTCCCTC 1140  
 QY 1141 ACGGAGTCCCATGACAGCAGCCCAAGGAAAGCCATCCCAAGAAAGATATATTGGT 1200  
 Db 1141 ACGGAGTCCCATGACAGCAGCCCAAGGAAAGCCATCCCAAGAAAGATATATTGGT 1200  
 QY 1201 ATCCGAATGATGACATCACTGATCAGCAAGAAAGCAAGAGCTGAAGACCGGACCGGAC 1260  
 Db 1201 ATCCGAATGATGACATCACTGATCAGCAAGAAAGCAAGAGCTGAAGACCGGACCGGAC 1260  
 QY 1261 TTCCGACAGCTGATCTCAGGAGCGTATATATTGAAGTAATTCCTGATACCCCGAGAAA 1320  
 Db 1261 TTCCGACAGCTGATCTCAGGAGCGTATATATTGAAGTAATTCCTGATACCCCGAGAAA 1320  
 QY 1321 GCTGGTGGTCTCAAGGAAAGCAAGTCAATATCAGCATCAATGGAAGTGGTGGTCTCC 1380  
 Db 1321 GCTGGTGGTCTCAAGGAAAGCAAGTCAATATCAGCATCAATGGAAGTGGTGGTCTCC 1380  
 QY 1381 GCCAATGATGTCAGCGAGCTCATTAAGAGGAAAGCAACCTGAAATGGTGGTCCGACG 1440  
 Db 1381 GCCAATGATGTCAGCGAGCTCATTAAGAGGAAAGCAACCTGAAATGGTGGTCCGACG 1440  
 QY 1441 GGTAAAGAGATATCATGATCAAGTATCCCGAAGAAATGACCCATAGGACAGGCA 1500  
 Db 1441 GGTAAAGAGATATCATGATCAAGTATCCCGAAGAAATGACCCATAGGACAGGCA 1500  
 QY 1501 TGAGCTGGACTTCATGTTCCCTCAAGACTCTCCCGTGGATGAGGATGAGGACTCTCG 1560  
 Db 1501 TGAGCTGGACTTCATGTTCCCTCAAGACTCTCCCGTGGATGAGGATGAGGACTCTCG 1560  
 QY 1561 GCTGCTGGAATAGGACACTCAAGCTTTTGTGATGTCATTTGTTTGTTCAGTGGAGCTC 1620  
 Db 1561 GCTGCTGGAATAGGACACTCAAGCTTTTGTGATGTCATTTTGTTCAGTGGAGCTC 1620  
 QY 1621 CTTGGCCCAAGAAATCTCTTTGATGTTTGAGGCAAAACAAATGTAATTTGAGATC 1680  
 Db 1621 CTTGGCCCAAGAAATCTCTTTGATGTTTGAGGCAAAACAAATGTAATTTGAGATC 1680  
 QY 1681 CGCAGGAGAGCTCTGCCCTTCTGTATGTCCTATGATGATGATGATGATGATGATGATG 1740  
 Db 1681 CGCAGGAGAGCTCTGCCCTTCTGTATGTCCTATGATGATGATGATGATGATGATGATG 1740  
 QY 1741 TTGGCCATCTCTGCTAGACAGTCAAGTATGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1800  
 Db 1741 TTGGCCATCTCTGCTAGACAGTCAAGTATGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1800  
 QY 1801 TAGTCCAACTAATGAGTCTGATACAAATCGTAGATAGAAGAGAGCCCGGAGCCAGGA 1860

Db 1801 TAGTCCAACTAATGAGTCTGATACAAATCGTAGATAGAAGAGAGCCCGGAGCCAGGA 1860  
 QY 1861 TGGGACTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1920  
 Db 1861 TGGGACTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1920  
 QY 1921 CCCGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1980  
 Db 1921 CCCGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1980  
 QY 1981 TTGAATTTGGGAGCAGCAGTACTCTGAGTTTGAGCTATTAAAGTACTTCTTACACA 2036  
 Db 1981 TTGAATTTGGGAGCAGCAGTACTCTGAGTTTGAGCTATTAAAGTACTTCTTACAAA 2036

RESULT 3

AAT75444  
 ID AAT75444 standard; cDNA; 2075 BP.

XX AAT75444;

DT 16-SEP-1997 (first entry)

XX cDNA encoding osteoblast like cell derived protein.

XX Osteoblast like cell; prevention; treatment; disease; contraceptive; ds.

XX Homo sapiens.

XX Key Location/Qualifiers

FT CDS 88..1530

FT /tag= a

FT sig\_peptide 88..168

FT /tag= b

FT mat\_peptide 169..1527

FT /tag= c

XX JP09107980-A.

XX 28-APR-1997.

XX 12-AUG-1996; 96JP-00231415.

XX 17-AUG-1995; 95JP-00233537.

XX (NIBS ) JAPAN TOBACCO INC.

XX WPI; 1997-292469/27.

XX P-PSDB; AAW22849.

XX DNA encoding osteoblast like cell derived protein - useful for treatment and prevention of various diseases and as contraceptive.

XX Claim 1; Page 21-23; 42pp; Japanese.

XX The present sequence encodes an osteoblast like cell derived protein, which may be used to prevent or treat various diseases, or as a contraceptive. mRNA was collected from an osteoblast like cell, and double stranded DNA synthesised. A 3'-directed cDNA library was prepared, and the base sequence of the DNA of each clone in the cDNA library determined. An osteoblast complete chain cDNA library was prepared, and the clone GS2422, which has a complete chain, isolated. GS2422 protein was expressed and purified, and an antibody against GS2422 protein was analysed. The GS2422 gene is expressed in various human organs

XX Sequence 2075 BP; 459 A; 592 C; 603 G; 421 T; 0 U; 0 Other;

Query Match 99.0%; Score 2034.4; DB 2; Length 2075;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2035; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CCGGCGCTTCCGCTGTCGCGCGCCACCGCGCGCCGCGCGAGTCGCGCATGCGAGTCCCG 60  
 Db 40 CCGGCGCTTCCGCTGTCGCGCGCCACCGCGCGCCGCGCGAGTCGCGCATGCGAGTCCCG 99  
 QY 61 CCGGCGCTTCCGCTGTCGCGCGCCACCGCGCGCCGCGCGAGTCGCGCATGCGAGTCCCG 120  
 Db 100 CCGGCGCTTCCGCTGTCGCGCGCCACCGCGCGCCGCGCGAGTCGCGCATGCGAGTCCCG 159  
 QY 121 TCCGCGCGCGCGCTGTCGCGCGCCACCGCGCGCCGCGCGAGTCGCGCATGCGAGTCCCG 180  
 Db 160 TCCGCGCGCGCGCTGTCGCGCGCCACCGCGCGCCGCGCGAGTCGCGCATGCGAGTCCCG 219  
 QY 181 CCGTCCCGCGCGCGCTGTCGCGCGCCACCGCGCGCCGCGCGAGTCGCGCATGCGAGTCCCG 240  
 Db 220 CCGTCCCGCGCGCGCTGTCGCGCGCCACCGCGCGCCGCGCGAGTCGCGCATGCGAGTCCCG 279  
 QY 241 TCCGAGTGTGCG 300  
 Db 280 TCCGAGTGTGCG 339  
 QY 301 GAGGCGCTGCG 360  
 Db 340 GAGGCGCTGCG 399  
 QY 361 GCGCAGCGCGCGCGCTGTCG 420  
 Db 400 GCGCAGCGCGCGCGCTGTCG 459  
 QY 421 ACCTAGCG 480  
 Db 460 ACCTAGCG 519  
 QY 481 CCGCGCGCTCATCGCTGCG 540  
 Db 520 CCGCGCGCTCATCGCTGCG 579  
 QY 541 TTGCGCGATAAATATACTTTATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600  
 Db 580 TTGCGCGATAAATATACTTTATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 639  
 QY 601 CATATCGAATTTGTTGCGAAGCTTCCGTTTCTAAACGAGAGGTGCGCGGTGCGTGTGCGG 660  
 Db 640 CATATCGAATTTGTTGCGAAGCTTCCGTTTCTAAACGAGAGGTGCGCGGTGCGTGTGCGG 699  
 QY 661 TCTGGGTTATTTGTCGGAAGATGGAATGATGATGATGATGATGATGATGATGATGATGATG 720  
 Db 700 TCTGGGTTATTTGTCGGAAGATGGAATGATGATGATGATGATGATGATGATGATGATGATG 759  
 QY 721 AAGCACCGGTGCAAGTTGAGCTGAAGAAAGCGTGCCTTACGAAGCGCAAAATCAAGGAT 780  
 Db 760 AAGCACCGGTGCAAGTTGAGCTGAAGAAAGCGTGCCTTACGAAGCGCAAAATCAAGGAT 819  
 QY 781 GTGGATGAGAAAGCAGACATCGCATCATCAAAATGACCAAGGCGCAAGCTGCTGTC 840  
 Db 820 GTGGATGAGAAAGCAGACATCGCATCATCAAAATGACCAAGGCGCAAGCTGCTGTC 879  
 QY 841 CTGCTGCTTGGCGCTCCTCAGAGCTCGCGCGGAGAGTTGCTGCTGCTGCTGCTGCTGCTG 900  
 Db 880 CTGCTGCTTGGCGCTCCTCAGAGCTCGCGCGGAGAGTTGCTGCTGCTGCTGCTGCTGCTG 939  
 QY 901 CCGTTTTCCTTTCAAAACACAGTCACACCGGAGTCGTGAGCACACCCAGCGAGGCGGC 960  
 Db 940 CCGTTTTCCTTTCAAAACACAGTCACACCGGAGTCGTGAGCACACCCAGCGAGGCGGC 999  
 QY 961 AAAGAGCTGGGCTCGCAACTCAGACATGCACTACATCCAGCGCGCGCGCATCATCAAC 1020  
 Db 1000 AAAGAGCTGGGCTCGCAACTCAGACATGCACTACATCCAGCGCGCGCGCATCATCAAC 1059  
 QY 1021 TATGGAACCTCGGAGCGCGCTTAGTAAACCTTGACCGGTGAAGTGAATTTGAACACT 1080  
 Db 1060 TATGGAACCTCGGAGCGCGCTTAGTAAACCTTGACCGGTGAAGTGAATTTGAACACT 1119  
 QY 1081 TTGAAGTGCACGCTGGAAATCTCTCTTTGCAATCCCATCTGATTAAGNTAAAGTTCCTC 1140

Db 1120 TTGAAGTGCACGCTGGAAATCTCTCTTTGCAATCCCATCTGATAAGATTTAAAGTTCCTC 1179  
 QY 1141 ACCGAGTCCCATGACCCGACGAGGCAAGGAAAGCCATCACCAGAAAGATATATTGGT 1200  
 Db 1180 ACCGAGTCCCATGACCCGACGAGGCAAGGAAAGCCATCACCAGAAAGATATATTGGT 1239  
 QY 1201 ATCCGAATGATGATCACTCAGCTCCAGCAAAAGCCAAAGAGCTGAAAGGACCGGAC 1260  
 Db 1240 ATCCGAATGATGATCACTCAGCTCCAGCAAAAGCCAAAGAGCTGAAAGGACCGGAC 1299  
 QY 1261 TTCCGAGACGTGATCTCAGGAGCGTATATAATTGAAGTAAATTCCTGATACCCAGAGAA 1320  
 Db 1300 TTCCGAGACGTGATCTCAGGAGCGTATATAATTGAAGTAAATTCCTGATACCCAGAGAA 1359  
 QY 1321 GCTGGTGTCTCAAGGAAAACGAGTCATATCAAGCATCAATGGAACAGTCCGTGCTCC 1380  
 Db 1360 GCTGGTGTCTCAAGGAAAACGAGTCATATCAAGCATCAATGGAACAGTCCGTGCTCC 1419  
 QY 1381 GCCAATGATCTCAGCGACGTCATTAAAGGGAAGAACCCCTGAACATGCTGCTCCAGG 1440  
 Db 1420 GCCAATGATCTCAGCGACGTCATTAAAGGGAAGAACCCCTGAACATGCTGCTCCAGG 1479  
 QY 1441 GGTAAATGAAGATATCATGATCACAGTGAATCCCGAAGAAATTGACCATAGGCAGAGCA 1500  
 Db 1480 GGTAAATGAAGATATCATGATCACAGTGAATCCCGAAGAAATTGACCATAGGCAGAGCA 1539  
 QY 1501 TGAGCTGGACTTCATGTTTCCCTCAAGACTCTCCGCTGGATGAGGATGAGGACTCTGG 1560  
 Db 1540 TGAGCTGGACTTCATGTTTCCCTCAAGACTCTCCGCTGGATGAGGATGAGGACTCTGG 1599  
 QY 1561 GCTGCTGGAATAGGACACTCAAGACTTTTGACTGCTGCAATTTGTTGTTGTTGTTGTTG 1620  
 Db 1600 GCTGCTGGAATAGGACACTCAAGACTTTTGACTGCTGCAATTTGTTGTTGTTGTTG 1659  
 QY 1621 CCTGGCCACAGAACTCTTCTGATAGTTTGCAGGCAAAACAAATGTAATGTTGAGATC 1680  
 Db 1660 CCTGGCCACAGAACTCTTCTGATAGTTTGCAGGCAAAACAAATGTAATGTTGAGATC 1719  
 QY 1681 CGCAGCAGAAAGCTCTGCTTCTGATCCTATGATGATGATGATGATGATGATGATGATGATG 1740  
 Db 1720 CGCAGCAGAAAGCTCTGCTTCTGATCCTATGATGATGATGATGATGATGATGATGATGATG 1779  
 QY 1741 TTGGGCGATTTCTGCTTAGACAGTCAGCATTTGCTCTCTCTTAACTGAGTCATCATCT 1800  
 Db 1780 TTGGGCGATTTCTGCTTAGACAGTCAGCATTTGCTCTCTCTTAACTGAGTCATCATCT 1839  
 QY 1801 TAGTCCAACTAAATGCAATGCGATGATCAATGCGTAGATAGAAAGCCCGAGGAGCCAGGA 1860  
 Db 1840 TAGTCCAACTAAATGCAATGCGATGATCAATGCGTAGATAGAAAGCCCGAGGAGCCAGGA 1899  
 QY 1861 TGGGACTGCTGCTGCTTCTGCTTCTCCTGATGAGCAACCCAAAGGTCAATGCAAGAG 1920  
 Db 1900 TGGGACTGCTGCTGCTTCTGCTTCTCCTGATGAGCAACCCAAAGGTCAATGCAAGAG 1959  
 QY 1921 CCCCGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1980  
 Db 1960 CCCCGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2019  
 QY 1981 TTGGAATTTGGAGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2036  
 Db 2020 TTGGAATTTGGAGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2075

RESULT 4  
 AAV29540  
 ID AAV29540 standard; cDNA; 2036 BP.  
 XX  
 AC AAV29540;  
 XX  
 DT 13-OCT-1998 (first entry)  
 XX  
 DE Homo sapiens D87258 sequence.



QY 1441 GGTAATGAAGATATCATGATCACAGTGTGATCCCGAAGAAATGACCCATAGCAGAGGCA 1500  
 Db 1441 GGTAATGAAGATATCATGATCACAGTGTGATCCCGAAGAAATGACCCATAGCAGAGGCA 1500  
 QY 1501 TGAGCTGGACTTCATGTTTCCCTCAAGACATCTCCCGTGGATGACGGATGAGACTCTGG 1560  
 Db 1501 TGAGCTGGACTTCATGTTTCCCTCAAGACATCTCCCGTGGATGACGGATGAGACTCTGG 1560  
 QY 1561 GCTGCTGGAATAGACACTCAAGACTTTTGTGATGCTGCTTGTGTTTGTTCAGTGGAGACTC 1620  
 Db 1561 GCTGCTGGAATAGACACTCAAGACTTTTGTGATGCTGCTTGTGTTTGTTCAGTGGAGACTC 1620  
 QY 1621 CTTGGCCCAACAGAAATCTCTTGTGATGTTTGCAGGCAAAACAAATGTAATGTTGCAGATC 1680  
 Db 1621 CTTGGCCCAACAGAAATCTCTTGTGATGTTTGCAGGCAAAACAAATGTAATGTTGCAGATC 1680  
 QY 1681 CGAGGCAAGACTCTGCCCTTCTGTATCTGTATGATGATGATGATGATGATGATGATGATG 1740  
 Db 1681 CGAGGCAAGACTCTGCCCTTCTGTATCTGTATGATGATGATGATGATGATGATGATGATG 1740  
 QY 1741 TTGGGCAATCTTCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1800  
 Db 1741 TTGGGCAATCTTCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1800  
 QY 1801 TAGTCCAACTAATGCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1860  
 Db 1801 TAGTCCAACTAATGCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1860  
 QY 1861 TGGGACTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1920  
 Db 1861 TGGGACTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1920  
 QY 1921 CCCGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGG 1980  
 Db 1921 CCCGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGG 1980  
 QY 1981 TTGGAATTTGGGAGCAGCATGATCTGAGTTTGGAGCTATTAAAGTACTTCTTACACA 2036  
 Db 1981 TTGGAATTTGGGAGCAGCATGATCTGAGTTTGGAGCTATTAAAGTACTTCTTACACA 2036

RESULT 5  
 AA004680  
 ID AA004680 standard; cdna; 2205 BP.

XX AC AA004680;

DT 20-JUL-1998 (first entry)

DE 15-JAN-1998 Human presenilin interacting protein gene mutTM1-TM2.

XX Presenilin interacting protein; mutTM1-TM2; serine protease;  
 KW heat shock protein; human; familial Alzheimer's disease; PAD; diagnosis;  
 KW therapy; ss.

XX OS Homo sapiens.

XX PN WO9801549-A2.

XX PD 15-JAN-1998.

XX PF 04-JUL-1997; 97WO-CA000475.

XX PR 05-JUL-1996; 96US-0021673P.

XX PR 12-JUL-1996; 96US-0021700P.

XX PR 08-NOV-1996; 96US-0029899P.

XX PR 02-JAN-1997; 97US-0034590P.

XX (UTOR ) UNIV TORONTO GOVERNING COUNCIL.

PA (HSCR-) HSC RES & DEV LP.

XX St George- Hyslop PH, Fraser PE, Rommens JM;

XX WPI; 1998-286355/25.  
 XX New isolated mutant presenilin-1 genes - useful for developing products  
 PT for use in detection, diagnosis and therapy of Alzheimer's disease and  
 PT for drug screening.  
 XX Claim 4; Page 222-223; 238pp; English.  
 XX This DNA sequence, mutTM1-TM2, codes for an unknown protein that  
 CC interacts with presenilin proteins (see AAW23964-67). To identify such  
 CC proteins (see also AAV04674-80), a yeast two-hybrid system was used to  
 CC screen a brain cDNA library for clones which interacted with functional  
 CC domains of the presenilins. The mutTM1-TM2 clone interacts with a mutant  
 CC presenilin-1 transmembrane 1-2 loop domain, and appears to correspond to  
 CC a known heat-shock serine protease gene. Methods for identifying  
 CC substances that affect the interaction of a presenilin-interacting  
 CC protein with presenilin are provided. Such substances may be useful in  
 CC the treatment of Alzheimer's disease. Transgenic animals carrying the  
 CC mutTM1-TM2 nucleic acid, pure preparations of the mutTM1-TM2 protein, and  
 CC antibodies that selectively bind mutTM1-TM2 protein are also claimed  
 XX  
 SQ Sequence 2205 BP; 525 A; 609 C; 614 G; 439 T; 0 U; 18 Other;  
 Query Match 97.9%; Score 2010.8; DB 2; Length 2205;  
 Best Local Similarity 98.6%; Pred. No. 0;  
 Matches 2026; Conservative 3; Mismatches 25; Indels 1; Gaps 1;  
 QY 1 CCGGCCCTCGCCCTGTCCGCGCCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 60  
 Db 76 CCGGCCCTCGCCCTGTCCGCGCCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 135  
 QY 61 CGCGCGCTCTTCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120  
 Db 136 CGCGCGCTCTTCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 195  
 QY 121 TCCCGGCGCGCGCGCTCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180  
 Db 196 TCCCGGCGCGCGCGCTCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 255  
 QY 181 CGCTGCG 239  
 Db 256 CGCTGCG 315  
 QY 240 CTGCGAGGTGTGCG 299  
 Db 316 CTGCGAGGTGTGCG 375  
 QY 300 CGAGGCGGTGCGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 359  
 Db 376 CGAGGCGGTGCGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 435  
 QY 360 CGCGCAGCGCGCGCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 419  
 Db 436 CGCGCAGCGCGCGCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 495  
 QY 420 CACCTAGCCCAACCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 479  
 Db 496 CACCTAGCCCAACCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 555  
 QY 480 GCCCGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 539  
 Db 556 GCCCGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 615  
 QY 540 TTTCGCCCATATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 599  
 Db 616 TTTCGCCCATATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 675  
 QY 600 TCATATCGAATTTGTTTCGAGAGCTTCGTTTCTTAAACAGAGAGTGGCGGTGGTGGTGGTGG 659  
 Db 676 TCATATCGAATTTGTTTCGAGAGCTTCGTTTCTTAAACAGAGAGTGGCGGTGGTGGTGGTGG 735  
 QY 660 GTCTGGGTTTATTGTTGTCGAGAGATGGAGTCTGTCGACAAATGCCACCGTGGTGGTGGTGG 719



1816	CTTGGGCCATTTTTCCTAGACAGTCAGCAATTGTNTCTCTTTAACTGAGTCATCATN	1871
1800	TTAGTCCAACTAATCAGTCGATACAATCGTAGATAGAAGAGCCCAACGGGAGCCAGG	1859
1876	TTAGTCCAACTAATCAGTCGATACAATCGTAGATAGAAGAGCCCAACGGGAGCCAGG	1935
1860	ATGGGACTGGTCGTGTGTGTCTTCTCCAAGTCAGCACCCCAAAGGTCAATGCACAGAG	1919
1936	ATGGGACTGGTGTGTGTGTCTTCTCCAAGTCAGCACCCCAAAGGTCAATGCACAGAG	1995
1920	ACCCCGGTGGTGAGCGCTGCTCTCAAACGGCCGAAGTGGCTCTTTTAGGAATCTC	1979
1996	ACCCCGGTGGTGAGCGCTGCTCTCAAACGGCCGAAGTGGCTCTTTTAGGAATNTN	2055
1980	TTTGGAAATTGGGAGCAGCAGATGACTCTGAGCTTTTGAGCTATTAAAGTACTTCTTACACATTG	2039
2056	TTTGGAAATTGGGAGCAGCAGATGAGTTGAGTTTGAGNTATTAAAGTANTTTTACACATTG	2115
2040	AAAAAAAAAAAAAAAAA	2054
2116	AAAAAAAAAAAAAAAAA	2130

RESULT 6  
AAF21747  
ID AAF21747 standard; DNA; 1855 BP.  
XX AAF21747;  
XX  
XX  
XX 27-MAR-2001 (first entry)  
XX  
XX Human breast and ovarian cancer associated antigen gene SEQ ID 134.  
XX Human; breast cancer; ovarian cancer; cytostatic; immunosuppressive;  
XX nontropic; neuroprotective; antiviral; antiallergic; hepatotropic;  
XX antidiabetic; antiinflammatory; antiulcer; vulnery; anticonvulsant;  
XX antibacterial; antifungal; antiparasitic; cardiac; immune disorder;  
XX Addison's disease; allergy; autoimmune haemolytic anaemia;  
XX autoimmune thyroiditis; diabetes mellitus; Crohn's disease;  
XX multiple sclerosis; rheumatoid arthritis; ulcerative colitis;  
XX cardiovascular disorder; wound healing; neurological disease; ds.  
XX Homo sapiens.  
XX  
XX WO200055173-A1.  
XX  
XX 21-SEP-2000.  
XX  
XX 08-MAR-2000; 2000WO-US005881.  
XX  
XX 12-MAR-1999; 99US-0124270P.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Rosen CA, Ruben SM;  
XX  
XX WPI; 2000-611515/58.  
XX  
XX P-PSDB; AAB58844.  
XX  
XX  
XX New human breast and ovarian cancer associated gene sequences and the  
XX polypeptides encoded by these genes, useful in the prevention, treatment  
XX and diagnosis of cancer, immune disorders, cardiovascular disorders and  
XX neurological diseases.  
XX  
XX Claim 1; Page 581; 1299pp; English.  
XX  
XX Sequences AAF21614 - AAF22031 represent DNA sequences encoding human  
XX proteins AAB58711 - AAB59128. The DNA and protein sequences are  
XX associated with breast and ovarian cancer. Included in the invention are  
XX sequences AAF22032 - AAF22040 and AAB59129 which are used in the  
XX isolation and characterisation of the DNA and protein sequences of the  
XX invention. The breast and ovarian cancer associated DNA, protein, agonist  
XX or antagonist sequences exhibit cytostatic; immunosuppressive; nontropic;  
XX or antagonist sequences exhibit cytostatic; immunosuppressive; nontropic;  
XX

CC	neuroprotective; antiviral; anti-allergic; hepatotropic; antidiabetic;	
CC	anti-inflammatory; antitumor; antileukemic; anticonvulsant; antibacterial;	
CC	antifungal; antiparasitic and cardant activity. The polynucleotide and	
CC	protein sequences are used in the diagnosis of cancer, particularly	
CC	breast and ovarian cancer. The nucleic acid sequences, proteins, agonists	
CC	and agonists may also be used in the diagnosis, prevention and treatment	
CC	of immune disorders e.g. Addison's disease, allergies, autoimmune	
CC	haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's	
CC	disease, multiple sclerosis, rheumatoid arthritis and ulcerative colitis;	
CC	cardiovascular disorders such as myocardial ischaemia; wound healing;	
CC	neurological diseases such as cerebral anoxia and epilepsy; and	
CC	infectious diseases	
XX		
SQ	Sequence 1855 BP; 457 A; 490 C; 515 G; 391 T; 0 U; 2 Other;	
	Query Match 86.7%; Score 1780; DB 3; Length 1855;	
	Best Local Similarity 99.9%; Pred. No. 0;	
	Matches 1802; Conservative 0; Mismatches 0; Indels 2; Gaps 2;	
Qy	251 GGGGCGCGCCGAGGGCGCGCGTGGGCGCTGCGAGGAGGCGCGTGGGCGAGGGGCTGC	310
Db	14 GGGGCGCGCCGAGGGCGCGCGTGGGCGCTGCGAGGAGGCGCGTGGGCGAGGGGCTGC	73
Qy	311 AGTGGTGGTGGCGCTTTCGGGGTCCGAGCCTTCGGCCACGCGTGGGCGCGCGCGCGAGGCGG	370
Db	74 AGTGGTGGTGGCGCTTTCGGGGTCCGAGCCTTCGGCCACGCGTGGGCGCGCGCGAGGCGG	133
Qy	371 GCCTCTGTGTGGCGCAGGAGCGCGGTGTGGCGAGGCGCGGAGCCCAACCTACGCCA	430
Db	134 GCCTCTGTGTGGCGCAGGAGCGCGGTGTGGCGAGGCGCGGAGCCCAACCTACGCCA	192
Qy	431 ACCTGTGCGAGGTCG	490
Db	193 ACCTGTGCGAGGTCG	252
Qy	491 TCGTCTGCGAGGTCG	550
Db	253 TCGTCTGCGAGGTCG	312
Qy	551 AATATAAATTTATCGCGGAGCGTGGTGGAGAGATCGCCCTCGCGTGGTTCATATCGAAT	610
Db	313 AATATAAATTTATCGCGGAGCGTGGTGGAGAGATCGCCCTCGCGTGGTTCATATCGAAT	372
Qy	611 TGTTTGGAGAGCTTCGTTTCTAAGCAGAGGTGGCGGTGGTGGTGGTGGTGGTGGTGGTGG	670
Db	373 TGTTTGGAGAGCTTCGTTTCTAAGCAGAGGTGGCGGTGGTGGTGGTGGTGGTGGTGGTGG	432
Qy	671 TTGTGTGCGAAGATGGACTGATCGTGACAAATGCGGAGGTCGCGCGCGCGCGCGCGCGG	730
Db	433 TTGTGTGCGAAGATGGACTGATCGTGACAAATGCGGAGGTCGCGCGCGCGCGCGCGCGG	492
Qy	731 TCAGAGTTCGAGTGAAGACGGTGGCGGCTTACGAGGCGGAGATCCCAAGTTGGCGGATA	790
Db	493 TCAGAGTTCGAGTGAAGACGGTGGCGGCTTACGAGGCGGAGATCCCAAGTTGGCGGATA	552
Qy	791 AAGCAGACATCGCACTCATCAAAATTTGACCAACCAAGGCAAGCTGCTGCTGCTGCTGCTG	850
Db	553 AAGCAGACATCGCACTCATCAAAATTTGACCAACCAAGGCAAGCTGCTGCTGCTGCTGCTG	612
Qy	851 GCGGCTCTCAGAGCTGCGCGCGGAGGTTGCTGGCTGCGCATCGGAGCGCGGTTTCCCG	910
Db	613 GCGGCTCTCAGAGCTGCGCGCGGAGGTTGCTGGCTGCGCATCGGAGCGCGGTTTCCCG	672
Qy	911 TTCAAACACAGTCACACCGGATCGTGAGCACCACCGCGAGGCGGCGGCAAGAGCTGG	970
Db	673 TTCAAACACAGTCACACCGGATCGTGAGCACCACCGCGAGGCGGCGGCAAGAGCTGG	732
Qy	971 GGCTCCGAGCTCAGACATCGATACACGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	1030
Db	733 GGCTCCGAGCTCAGACATCGATACACGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	792
Qy	1031 CGGGAGGCGCGGTTAGTAAACCTGGAGCGGTGAAGTGAATTAACACTTTGAAAGTGA	1090

RESULT 7  
ADB62641

Db	793 CGGGAGGCGCGTGTAGTAAACCTGGAGCGTGAAGTGAATTAACACTTTTGAAGTGA	852
Qy	1091 CAGCTGGAGTCTCTTGGCAATCCCATCTGATAGATTAAAGTCTCTCTCAGGAGTCC	1150
Db	853 CAGCTGGAGTCTCTTGGCAATCCCATCTGATAGATTAAAGTCTCTCTCAGGAGTCC	912
Qy	1151 ATGACCGACAGGCGCAAGGAAAGCCATCACCAAGAAAGTATATTTGGTATCCGAATGA	1210
Db	913 ATGACCGACAGGCGCAAGGAAAGCCATCACCAAGAAAGTATATTTGGTATCCGAATGA	972
Qy	1211 TGCTACTCAGTCCAGCAAGCCAAAGAGCTGAAGGACCGGACCGGAGCTTCCGAGCG	1270
Db	973 TGCTACTCAGTCCAGCAAGCCAAAGAGCTGAAGGACCGGACCGGAGCTTCCGAGCG	1032
Qy	1271 TGATCTCAGAGCGTATATAATTTGAAGTAATCTCTGATACCCAGCAGAGCTGGTGTG	1330
Db	1033 TGATCTCAGAGCGTATATAATTTGAAGTAATCTCTGATACCCAGCAGAGCTGGTGTG	1092
Qy	1331 TCAAGGAAACACGCTCATTAATCAGCATCAATGAGACAGTCCGTGGTCTCCGCAATGATG	1390
Db	1093 TCAAGGAAACACGCTCATTAATCAGCATCAATGAGACAGTCCGTGGTCTCCGCAATGATG	1152
Qy	1391 TCAGGACGCTCATTAATTAAGGAAAGCACCTGTAACATGGTGGTCCGACGGGTAATGAAG	1450
Db	1153 TCAGGACGCTCATTAATTAAGGAAAGCACCTGTAACATGGTGGTCCGCA- GGGTAAATGAAG	1211
Qy	1451 ATATCATGATCAAGTATTTCCCGAGAAATTTGACCCATAGGACAGGAGCATGAGCTGGAC	1510
Db	1212 ATATCATGATCAAGTATTTCCCGAGAAATTTGACCCATAGGACAGGAGCATGAGCTGGAC	1271
Qy	1511 TTATGTTTTCCCTCAAAGACTCTCCCGTGGATGACGAGTGAAGGACTCTGGGCTGCTGGAA	1570
Db	1272 TTATGTTTTCCCTCAAAGACTCTCCCGTGGATGACGAGTGAAGGACTCTGGGCTGCTGGAA	1331
Qy	1571 TAGGACACTCAAGACTTTTGGCTGCCATTTTGTGTTTGTTCAGTGAGAGCTCCCTGGCCAAAC	1630
Db	1332 TAGGACACTCAAGACTTTTGGCTGCCATTTTGTGTTTGTTCAGTGAGAGCTCCCTGGCCAAAC	1391
Qy	1631 AGAATCCTTCTTGATAGTTTGGAGGCAAAACAAATGTAATGTTGAGATCCCGACGACAG	1690
Db	1392 AGAATCCTTCTTGATAGTTTGGAGGCAAAACAAATGTAATGTTGAGATCCCGACGACAG	1451
Qy	1691 AGCTCTGCCCTTCTGATCTCTATGATGACGAGTGTGCTTTTCTTGGCAGCTTGGGCAAT	1750
Db	1452 AGCTCTGCCCTTCTGATCTCTATGATGACGAGTGTGCTTTTCTTGGCAGCTTGGGCAAT	1511
Qy	1751 CTTGCTTAGACAGTCAAGATTTGCTCTCTCTCTTAACTGAGTCACTCTTAGTCCAAT	1810
Db	1512 CTTGCTTAGACAGTCAAGATTTGCTCTCTCTCTTAACTGAGTCACTCTTAGTCCAAT	1571
Qy	1811 AATGCACTCGATCAATCGTAGATAGAAAGCCCAAGGAGCCAGGATGGGACTGGT	1870
Db	1572 AATGCACTCGATCAATCGTAGATAGAAAGCCCAAGGAGCCAGGATGGGACTGGT	1631
Qy	1871 CGTGTGTTGCTTTTCTCAGAGTCAGACCCAAAGGTCATGACAGAGACCCCGGTTGG	1930
Db	1632 CGTGTGTTGCTTTTCTCAGAGTCAGACCCAAAGGTCATGACAGAGACCCCGGTTGG	1691
Qy	1931 GTGAGCGCTGGCTTCTCAAAAGCGGAGGTTGCTCTTTTAGGAACTCTTTTGGAAATGG	1990
Db	1692 GTGAGCGCTGGCTTCTCAAAAGCGGAGGTTGCTCTTTTAGGAACTCTTTTGGAAATGG	1751
Qy	1991 GAGCAGATGACTCTGAGTTTGAAGTAAAGTACTTCTTACACATTTGAAAGGAAAAA	2050
Db	1752 GAGCAGATGACTCTGAGTTTGAAGTAAAGTACTTCTTACACATTTGAAAGGAAAAA	
Qy	2051 AAAA 2054	
Db	1812 AAAA 1815	







Db 241 AGCCAAATCAAGATGTGGATGAGAAAGCAGACATCGCACTCATCAAAATTTGACCACCA 300  
 QY 825 GGGCAAGCTCCTGTCTCTGCTGCTTGGCCGCTCTCTCAGAGCTGCGCGCGGAGAGTTGCT 884  
 Db 301 GGGCAAGCTCCTGTCTCTGCTGCTTGGCCGCTCTCTCAGAGCTGCGCGCGGAGAGTTGCT 360  
 QY 885 GGTCCGCATCGGAGCCGTTTCCCTTCAAAACACAGTCCACCGGATCGTGAGCAC 944  
 Db 361 GGTCCGCATCGGAGCCGTTTCCCTTCAAAACACAGTCCACCGGATCGTGAGCAC 420  
 QY 945 CACCCAGCGAGCGGCAAGAGCTGGGCTCGCAACTCAGACATGGAATACATCCAGAC 1004  
 Db 421 CACCCAGCGAGCGGCAAGAGCTGGGCTCGCAACTCAGACATGGAATACATCCAGAC 480  
 QY 1005 CGAGCCATCATCAACTATGG--AACTCGGAGGCGCGTTAGTAACTCGGAGCGTGA 1062  
 Db 481 CGAGCCATCATCAACTATGGAACTCCGGAGGCGCGTTAGTAACTCGGAGCGTGA 540  
 QY 1063 GTGATTGGAATTAAACACTTTGAAAGTGAAGTGGAAATCTCTTTGCAATCCCATCTGAT 1122  
 Db 541 GTGATTGGAATTAAACACTTTGAAAGTGAAGTGGAAATCTCTTTGCAATCCCATCTGAT 600  
 QY 1123 AGATTAAAGTTCTCCAGGAGTCCATGACGAGCGGCAAGGAAAGCCATCAC 1182  
 Db 601 AGATTAAAGTTCTCCAGGAGTCCATGACGAGCGGCAAGGAAAGCCATCAC 660  
 QY 1183 AAGAAGATATATTGGTATCCGATGATGTCACCTCAGCTCCAGCAAGCCAAAGAGCTG 1242  
 Db 661 AAGAAGATATATTGGTATCCGATGATGTCACCTCAGCTCCAGCAAGCCAAAGAGCTG 720  
 QY 1243 AAGACCGGACCGGACTTCCAGAGCTGATCTCAGAGCGTATATATTGAAGTAAT 1302  
 Db 721 AAGACCGGACCGGACTTCCAGAGCTGATCTCAGAGCGTATATATTGAAGTAAT 780  
 QY 1303 CTTGATACCCAGCAGAGCTGGT--GGTCTCAAGGAAACAGCTCATATCAGCATCAA 1361  
 Db 781 CTTGATACCCAGCAGAGCTGGTGGTCTCAAGGAAACAGCTCATATCAGCATCAA 840  
 QY 1362 TGACAGCTCGTGCTCCGCAATGATGTGAGCAGC--TCATTAAGGAAAGCAGCC 1420  
 Db 841 TGACAGCTCGTGCTCCGCAATGATGTGAGCAGCTCATTAAGGAAAGCAGCC 900  
 QY 1421 TGAACATGGTGGTCCGAGGGGTAAAGATATCATGATCAGATGATTTCCCGAAGAA 1480  
 Db 901 TGAACATGGTGGTCCGAGGGGTAAAGATATCATGATCAGATGATTTCCCGAAGAA 960  
 QY 1481 TTGACCCATAGGAGCATGAGCTGAGCTCATGTTTCCCTCAAAGACTCTCCCGT-G 1539  
 Db 961 TTGACCCATAGGAGCATGAGCTGAGCTCATGTTTCCCTCAAAGACTCTCCCGTGG 1020  
 QY 1540 GATGACGGATGAGACTCTGGGCTGCTGGAATAGGACACTCAAGACTTTTGACTGCCATT 1599  
 Db 1021 GATGACGGATGAGACTCTGGGCTGCTGGAATAGGACACTCAAGACTTTTGACTGCCATT 1080  
 QY 1600 TTGTTTCTCAGTGGAGACTCCCTGGCACAAGAAATCTTTGATGATTTGAGTTCAGGCAAA 1659  
 Db 1081 TTGTTTCTCAGTGGAGACTCCCTGGCACAAGAAATCTTTGATGATTTGAGTTCAGGCAAA 1140  
 QY 1660 ACAAATGATTTGACATCGCAGGAGGAGCTG--CCCTTCTGATCTCTATGATG 1718  
 Db 1141 ACAAATGATTTGACATCGCAGGAGGAGCTCTGCCCCTTCTGATCTCTATGATG 1200  
 QY 1719 CAGTGTCTTTTCTTTCGACCTGGGCAATTCCTTGTGAGACAGTTCAGCATTTGCTCC 1778  
 Db 1201 CAGTGTCTTTTCTTTCGACCTGGGCAATTCCTTGTGAGACAGTTCAGCATTTGCTCC 1260  
 QY 1779 TCCTTTAAGTGAATCATCTTATGCAACTAATGAGCTGATCAATG--GATGATAG 1837  
 Db 1261 TCCTTTAAGTGAATCATCTTATGCAACTAATGAGCTGATCAATG--GATGATAG 1320  
 QY 1838 AAGAAGCCCAAGGAGCAGGATGGAGCTGCTGTTGCTTTTCTCAAGTCAGC 1897

Db 1321 AAGAAGCCCAAGGAGCAGGATGGAGTGGTCTGTTGTTGCTTTTCTCAAGTCAGC 1380  
 QY 1898 ACCCAAAGGTCAATGACAGAGACCCCGGTGGGTGAGCGTGGCTTCTCAAAAGGCGCA 1957  
 Db 1381 ACCCAAAGGTCAATGACAGAGACCCCGGTGGGTGAGCGTGGCTTCTCAAAAGGCGCA 1440  
 QY 1958 AGTTGCTCTTTTGGAGTCTCTTTTGGAAATTCGGAGCAGATGACTCTGAGTTTGAAGCTA 2017  
 Db 1441 AGTTGCTCTTTTGGAGTCTCTTTTGGAAATTCGGAGCAGATGACTCTGAGTTTGAAGCTA 1500  
 QY 2018 TTAAGTACTTCTTACACATTTGAAAAA 2051  
 Db 1501 TTAAGTACTTCTTACACATTTGAAAAA 1534

RESULT 10  
 AAD44188  
 ID AAD44188 standard; cDNA; 1440 BP.  
 XX  
 AC AAD44188;  
 XX  
 DT 13-DEC-2002 (first entry)  
 XX  
 DE Human TANGO 219 cDNA.  
 XX  
 KW TANGO protein; vaccine; chromosomal mapping; antisense gene therapy;  
 KW forensic biology; predictive medicine; cytostatic; pharmacogenomic;  
 KW tissue typing; human; chromosome 5; gene; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US2002055139-A1.  
 XX  
 PD 09-MAY-2002.  
 XX  
 PF 01-MAR-2001; 2001US-00796858.  
 XX  
 PR 30-DEC-1998; 98US-00223094.  
 PR 30-DEC-1998; 98US-00223546.  
 PR 30-DEC-1998; 98US-00224246.  
 PR 14-MAY-1999; 99US-00312359.  
 PR 18-JUN-1999; 99US-00336536.  
 PR 29-JUN-1999; 99US-00342687.  
 PR 30-JUL-1999; 99US-00365164.  
 PR 20-SEP-1999; 99US-00399723.  
 PR 21-DEC-1999; 99US-00471179.  
 PR 29-DEC-1999; 99US-00474071.  
 PR 29-DEC-1999; 99US-00474072.  
 PR 15-MAY-2000; 2000US-00572002.  
 PR 19-JUN-2000; 2000US-00597993.  
 PR 22-JUN-2000; 2000US-00599596.  
 PR 29-JUN-2000; 2000US-00606565.  
 PR 31-JUL-2000; 2000US-00630334.  
 PR 20-SEP-2000; 2000US-00655666.  
 XX  
 (HOLT/) HOLTZMAN D A.  
 PA (SHAR/) SHARP J D.  
 PA (LEIB/) LEIBY K R.  
 PA (BOSS/) BOSSONE S.  
 PA (PANY/) PAN Y.  
 PA (BARN/) BARNES T M.  
 PA (PRAS/) PRASER C C.  
 PA (WRIG/) WRIGHTON N.  
 PA (MYER/) MYERS P S.  
 PA (KING/) KINGSBURY G.  
 XX  
 PI Holtzman DA, Sharp JD, Leiby KR, Bossone S, Pan Y, Barnes TM;  
 PI Fraser CC, Wrighton N, Myers PS, Kingsbury G;  
 XX  
 DR WPI; 2002-453953/48.  
 XX  
 PT New isolated nucleic acid encoding a TANGO polypeptide, for use as a  
 PT modulating agent for regulating cellular processes and for use in a





```
FT      old_sequence      /*tag= c      replace(12, G)
FT      old_sequence      /*tag= d      replace(12, G)
FT      old_sequence      /*tag= e      replace(13, C)
FT      old_sequence      /*tag= f      replace(15, C)
FT      old_sequence      /*tag= g      replace(18, C)
FT      old_sequence      /*tag= h      replace(21, T)
FT      old_sequence      /*tag= i      replace(22, C)
FT      old_sequence      /*tag= j      replace(24, T)
FT      old_sequence      /*tag= k      replace(25, C)
FT      old_sequence      /*tag= l      replace(27, C)
FT      old_sequence      /*tag= m      replace(30, G)
FT      old_sequence      /*tag= n      replace(33, G)
FT      old_sequence      /*tag= o      replace(37, C)
FT      old_sequence      /*tag= p      replace(39, G)
FT      old_sequence      /*tag= q      replace(40, C)
FT      old_sequence      /*tag= r      replace(42, G)
FT      old_sequence      /*tag= s      replace(48, G)
FT      old_sequence      /*tag= t      replace(51, G)
FT      old_sequence      /*tag= u      replace(54, G)
FT      old_sequence      /*tag= v      replace(57, C)
FT      old_sequence      /*tag= w      replace(60, C)
FT      old_sequence      /*tag= x      replace(63, G)
FT      old_sequence      /*tag= y      replace(66, G)
FT      old_sequence      /*tag= z      replace(69, G)
FT      old_sequence      /*tag= aa     replace(72, G)
FT      old_sequence      /*tag= ab     replace(75, C)
FT      old_sequence      /*tag= ac     replace(78, G)
FT      old_sequence      /*tag= ad     replace(81, C)
FT      old_sequence      /*tag= ae     replace(84, C)
FT      old_sequence      /*tag= af     replace(85, C)
FT      old_sequence      /*tag= ag     replace(87, C)
FT      old_sequence      /*tag= ah     replace(90, G)
FT      old_sequence      /*tag= ai     replace(93, G)
FT      old_sequence      /*tag= aj     replace(96, T)
FT      old_sequence      /*tag= ak     replace(99, G)
FT      old_sequence      /*tag= al     replace(102, C)
FT      old_sequence      /*tag= am     replace(105, C)
FT      old_sequence      /*tag= an     replace(108, G)
FT      old_sequence      /*tag= ao     replace(111, C)
FT      old_sequence      /*tag= ap     replace(114, A)
FT      old_sequence      /*tag= aq     replace(117, C)
FT      old_sequence      /*tag= ar     replace(118, C)
FT      old_sequence      /*tag= as     replace(120, A)
FT      old_sequence      /*tag= at     replace(123, C)
FT      old_sequence      /*tag= au     replace(129, G)
FT      old_sequence      /*tag= av     replace(132, G)
FT      old_sequence      /*tag= aw     replace(133, C)
FT      old_sequence      /*tag= ax     replace(135, C)
FT      old_sequence      /*tag= ay     replace(138, C)
FT      old_sequence      /*tag= az     replace(141, G)
FT      old_sequence      /*tag= ba     replace(144, G)
FT      old_sequence      /*tag= bb     replace(147, G)
FT      old_sequence      /*tag= bc     replace(150, G)
FT      old_sequence      /*tag= bd     replace(153, G)
FT      old_sequence      /*tag= be     replace(156, C)
FT      old_sequence      /*tag= bf     replace(162, G)
FT      old_sequence      /*tag= bg     replace(165, C)
FT      old_sequence      /*tag= bh     replace(168, C)
FT      old_sequence      /*tag= bi     replace(169, C)
FT      old_sequence      /*tag= bj     replace(171, G)
FT      old_sequence      /*tag= bk     replace(174, C)
FT      old_sequence      /*tag= bl     replace(175, C)
FT      old_sequence      /*tag= bm     replace(177, G)
FT      old_sequence      /*tag= bn     replace(180, C)
FT      old_sequence      /*tag= bo     replace(183, G)
FT      old_sequence      /*tag= bp     replace(189, C)
FT      old_sequence      /*tag= bq     replace(192, C)
FT      old_sequence      /*tag= br     replace(198, G)
FT      old_sequence      /*tag= bs     replace(201, G)
FT      old_sequence      /*tag= bt     replace(207, C)
FT      old_sequence      /*tag= bu     replace(210, G)
FT      old_sequence      /*tag= bv     replace(213, C)
FT      old_sequence      /*tag= bw     replace(216, G)
FT      old_sequence      /*tag= bx     replace(219, C)
```



KW Serine protease; ASP05; Insulin-like Growth Factor Binding Protein;  
 XX IGFBP; antiproliferative; anti-angiogenic; cancer; ss.  
 OS Synthetic.  
 OS Homo sapiens.  
 XX Key Location/Qualifiers  
 FT CDS 1..1185  
 FT /\*tag= a  
 FT /product= "C-terminal end of serine protease (ASP05) with  
 FT Histidine tag"  
 XX W0955885-A2.  
 XX 04-NOV-1999.  
 XX 28-APR-1999; 99WO-US009224.  
 XX 28-APR-1998; 98US-0083321P.  
 XX 27-APR-1999; 99US-00300621.  
 XX (AXYS-) AXYS PHARM INC.  
 XX Smeekens SP, Hou J;  
 XX WPI; 2000-072182/06.  
 XX P-PSDB; AAY15193.  
 XX New recombinant nucleic acid, with reduced GC content, encoding the  
 PT serine protease ASP05, used for treating disorders involving insulin-like  
 PT growth factor, e.g. cancer.  
 XX Disclosure; Fig 6; 70pp; English.  
 XX The present sequence is a cDNA encoding the C-terminal domain of serine  
 CC protease, ASP05 with a histidine tag. It can be easily purified by  
 CC affinity purification using an anti-tag antibody. ASP05 specifically  
 CC cleaves insulin-like Growth Factor Binding protein (IGFBP) and shows  
 XX anticancer, antiproliferative and anti-angiogenic activity  
 XX  
 SQ Sequence 1185 BP; 338 A; 291 C; 307 G; 249 T; 0 U; 0 Other;  
 Query Match 50.8%; Score 1043.8; DB 3; Length 1185;  
 Best Local Similarity 95.4%; Pred. No. 5.7e-195;  
 Matches 1075; Conservative 0; Mismatches 52; Indels 0; Gaps 0;  
 362 CGCAGCGCGCTGTGTGTGCGCCAGCAGCGCGGTGCGCGCAGCGCAACA 421  
 35 CCCAAGCTGGTTATGTGTATGCGCGAGTTGAGAACCAAGTATGTGCTCTGATGCAATA 94  
 422 CCTACGCCAACCTGCGCAGCTGCGCGCGCGCGCTCCGAGAGCTGCCACCGC 481  
 95 CATACGCAAACTTATGCCAATTGAGAGCTGCTTTAGACGTAGTGAAGACTACATAGAC 154  
 482 CGCGCGTCATGCTCTGCGAGCGCGAGCTGCGGCAAGGCGAGGAGATCCCAACAGTT 541  
 155 CGCTGTATTAGTCTGCAACGGGAGCGCTGCGGCAAGGCGAGGAGATCCCAACAGTT 214  
 542 TGGCGCAATAATAACTTTATCGCGAGCGTGTGAGAGATCGCCCTGCGGTGTC 601  
 215 TGGCGCAATAATAACTTTATCGCGAGCGTGTGAGAGATCGCCCTGCGGTGTC 274  
 602 ATATCGAATTGTTTCGGAAGCTTCGTTTCTTAAACGAGAGGTGCGGTGAGTGGGT 661  
 275 ATATCGAATTGTTTCGGAAGCTTCGTTTCTTAAACGAGAGGTGCGGTGAGTGGGT 334  
 662 CTGGTTTATGTCGGAAGATGACTGATGCTGACAAATGCCACCTGCTGACCAACA 721  
 335 CTGGTTTATGTCGGAAGATGACTGATGCTGACAAATGCCACCTGCTGACCAACA 394  
 722 AGCACCGGTCGGAAGTGTGAGTGAAGACGCGTCCACTTACGAAGCCAAATCAAGGATG 781  
 395 AGCACCGGTCGGAAGTGTGAGTGAAGACGCGTCCACTTACGAAGCCAAATCAAGGATG 454

QY 782 TGGATGAGAAAGCAGACATCGCACTCATCAAAATTTGACCACGAGGCAAGCTCCTGTCC 841  
 DB 455 TGGATGAGAAAGCAGACATCGCACTCATCAAAATTTGACCACGAGGCAAGCTCCTGTCC 514  
 QY 842 TGCTGCTTGGCGGCTCCTCAGAGCTGCGCGGGAGAGTTCTGTGTCGCCATCGGAAGCC 901  
 DB 515 TGCTGCTTGGCGGCTCCTCAGAGCTGCGCGGGAGAGTTCTGTGTCGCCATCGGAAGCC 574  
 QY 902 CGTTTTCCCTTCAAAACACAGTCACACCGGATCGTGAAGCACCACCGAGGCGGCA 961  
 DB 575 CGTTTTCCCTTCAAAACACAGTCACACCGGATCGTGAAGCACCACCGAGGCGGCA 634  
 QY 962 AAGAGCTGGGCTCGGCACTCAGACATGGAATCATCCAGACCGGAGCCATCATCACT 1021  
 DB 635 AAGAGCTGGGCTCGGCACTCAGACATGGAATCATCCAGACCGGAGCCATCATCACT 694  
 QY 1022 ATGGAACCTCGGAGGCGGCTTAGTAAACCTTGAGCGGTGAAGTGATTTGAATTAACACT 1081  
 DB 695 ATGGAACCTCGGAGGCGGCTTAGTAAACCTTGAGCGGTGAAGTGATTTGAATTAACACT 754  
 QY 1082 TGAAGTGACAGCTGGATCTCTTTCATCCCTCATCTGATGAATTAAGTAAAGTCTCTCA 1141  
 DB 755 TGAAGTGACAGCTGGATCTCTTTCATCCCTCATCTGATGAATTAAGTAAAGTCTCTCA 814  
 QY 1142 CGGAGTCCCATGACCGACAGCGCCAAAGGAAAGCCATCACCAAGAAAGATATATTGGTA 1201  
 DB 815 CGGAGTCCCATGACCGACAGCGCCAAAGGAAAGCCATCACCAAGAAAGATATATTGGTA 874  
 QY 1202 TCCGAATGATGTCATCTCAGTCCAGCAAGGCAAGAGCTGAAGGACCGGACCGGACT 1261  
 DB 875 TCCGAATGATGTCATCTCAGTCCAGCAAGGCAAGAGCTGAAGGACCGGACCGGACT 934  
 QY 1262 TCCCAGACGTGATCTCAGGAGCGGTATATAATTGAAGTAAATTCCTGATACCCAGCAGAAG 1321  
 DB 935 TCCCAGACGTGATCTCAGGAGCGGTATATAATTGAAGTAAATTCCTGATACCCAGCAGAAG 994  
 QY 1322 CTGGTGTCTCAAGGAAACGACGTGATATCAGCATCATGAGCAGTCTGCTGCTCCG 1381  
 DB 995 CTGGTGTCTCAAGGAAACGACGTGATATCAGCATCATGAGCAGTCTGCTGCTCCG 1054  
 QY 1382 CCAATGATGTCAGCGACGTCTATTAAAGGAAAGCACCCTGAAACATGGTGGTCCGAGGG 1441  
 DB 1055 CCAATGATGTCAGCGACGTCTATTAAAGGAAAGCACCCTGAAACATGGTGGTCCGAGGG 1114  
 QY 1442 GTAATGAAGATATCATGATCATCAGTGATTCGGAAGAAATTCACCCA 1488  
 DB 1115 GTAATGAAGATATCATGATCATCAGTGATTCGGAAGAAATTCACCCA 1161

RESULT 13  
 ABX77560  
 ID ABX77560 standard; cDNA; 1389 BP.  
 XX  
 AC ABX77560;  
 XX  
 DT 09-APR-2003 (first entry)  
 XX  
 DE Differentially expressed breast cancer associated cDNA #55.  
 XX Breast cancer; differential gene expression; BC-cDNA;  
 XX breast cancer diagnosis; breast cancer monitoring;  
 XX breast cancer treatment; breast cancer staging; Gene; ss.  
 OS Homo sapiens.  
 XX  
 PN US2002156263-A1.  
 XX  
 PD 24-OCT-2002.  
 XX  
 PF 04-OCT-2001; 2001US-00974298.  
 XX  
 PR 05-OCT-2000; 2000US-0238331P.

XX PA (CHEN/) CHEN H.  
 XX PI Chen H;  
 XX DR WPI; 2003-182653/18.  
 XX PT New cDNAs, which are differentially expressed in (metastatic) breast  
 PT cancer useful for diagnosing or staging, breast cancer, or for monitoring  
 PT the treatment of breast cancer in an individual.  
 XX Claim 1; SEQ ID NO 67; 30pp; English.  
 XX The invention describes a combination of cDNAs (designated BC-cDNAs),  
 CC which are differentially expressed in breast cancer. The combination  
 CC includes 152 cDNA sequences, or their complements. The protein encoded by  
 CC any of these BC-cDNAs is useful for screening several molecules or  
 CC compounds to identify at least one ligand that specifically binds the  
 CC protein, producing or preparing polyclonal or monoclonal antibodies, or  
 CC purifying antibodies from a sample. The antibodies, which specifically  
 CC bind the protein differentially expressed in breast cancer is useful for  
 CC detecting the expression of a protein in a sample. The BC-cDNAs are also  
 CC useful for diagnosing, monitoring the treatment of, or staging, breast  
 CC cancer. This sequence represents a differentially expressed breast cancer  
 CC associated cDNA. Note: The sequence data for this patent did not form  
 CC part of the printed specification, but was obtained in electronic format  
 CC directly from the US patent office at  
 CC seqdata.uspto.gov/sequence.html?DocID=20020156263  
 XX SQ Sequence 1389 BP; 364 A; 324 C; 357 G; 343 T; 0 U; 1 Other;  
 Query Match 37.0%; Score 759.8; DB 7; Length 1389;  
 Best Local Similarity 99.5%; Pred. No. 2.5e-139;  
 Matches 772; Conservative 0; Mismatches 3; Indels 1; Gaps 1;  
 QY 1219 ACCTCAGCAAGCAAGAGCTGAAGACCGGACCGGACTTCCACAGCTGATCTCA 1278  
 DB 615 ACATCCAGCAAGCAAGAGCTGAAGACCGGACCGGACTTCCACAGCTGATCTCA 674  
 QY 1279 GGACGCTATATAATTGAAGTAACTTCTGATATCCCGACGAGAGCTGGTGTCTCAAGGAA 1338  
 DB 675 GGACGCTATATAATTGAAGTAACTTCTGATATCCCGACGAGAGCTGGTGTCTCAAGGAA 734  
 QY 1339 AACGAGCTATATCAGCATCATGAGACATCGGTGGTCTCGCCATGATGTCAGGAC 1398  
 DB 735 AACGAGCTATATCAGCATCATGAGACATCGGTGGTCTCGCCATGATGTCAGGAC 794  
 QY 1399 GTCAATTAAGGGAAGCAACCTGACATGGTGGTCCGACGGGTAAATGAAGATATCATG 1458  
 DB 795 GTCAATTAAGGGAAGCAACCTGACATGGTGGTCCGACGGGTAAATGAAGATATCATG 854  
 QY 1459 ATCAGATGATTCGGAAGAAATGACCATAGGACGAGGATGAGCTGGACTTCATGTT 1518  
 DB 855 ATCAGATGATTCGGAAGAAATGACCATAGGACGAGGATGAGCTGGACTTCATGTT 914  
 QY 1519 TCCCTCAAGAGCTCTCCCGTGATGACGATGAGGACTCTGGGCTGCTGGAATAGGACAC 1578  
 DB 915 TCTCTCAAGAGCTCTCCCGTGATGACGATGAGGACTCTGGGCTGCTGGAATAGGACAC 974  
 QY 1579 TCAGACTTTGACGCGCAATTTGTTTGTTCAGTGCAGACTCCCTGCCACAGATCCT 1638  
 DB 975 TCAGACTTTGACGCGCAATTTGTTTGTTCAGTGCAGACTCCCTGCCACAGATCCT 1034  
 QY 1639 TCTTGATAGTTGACGCAAAACAAATGTAATGTTGCAGATCCGACGAGGAGAGCTCTGC 1698  
 DB 1035 TCTTGATAG-ITGACGCAAAACAAATGTAATGTTGCAGATCCGACGAGGAGAGCTCTGC 1093  
 QY 1699 CTTCTGTATCTGATGAGTGTGCTTTCTTTCAGCTTGGCCATCTTCGTTA 1758  
 DB 1094 CTTCTGTATCTGATGAGTGTGCTTTCTTTCAGCTTGGCCATCTTCGTTA 1153  
 QY 1759 CACAGCTGACATTTGCTCCTCTTAACTGAGTATCATCTTAGTCCCACTAATGACGT 1818

DB 1154 GACAGTCAGCATTTGTCTCCTCTTAACTGAGTCATCATCTTAGTCCCACTAATGACGT 1213  
 QY 1819 CGATACAAATGCTAGATAGAAAGAGCCCAACCGGAGCCAGGATGGACTGGTGTGTTG 1878  
 DB 1214 CGATACAAATGCTAGATAGAAAGAGCCCAACCGGAGCCAGGATGGACTGGTGTGTTG 1273  
 QY 1879 TGCCTTTCTCCAAAGTCAGCAGCCCAAGGTCATGACACAGAGACCCCGGGTGGGTGAGCGC 1938  
 DB 1274 TGCCTTTCTCCAAAGTCAGCAGCCCAAGGTCATGACACAGAGACCCCGGGTGGGTGAGCGC 1333  
 QY 1939 TGCCTTTCTCCAAAGTCAGCAGCCCAAGTTCCTTTAGGAATCTCTTGAATTTGGAGC 1994  
 DB 1334 TGCCTTTCTCCAAAGTCAGCAGCCCAAGTTCCTTTAGGAATCTCTTGAATTTGGAGC 1389  
 RESULT 14  
 ID ABQ59513/c  
 ID ABQ59513 standard; cDNA; 669 BP.  
 AC ABQ59513;  
 DT 02-AUG-2002 (first entry)  
 DE Human colon cancer related nucleotide sequence SEQ ID NO:3208.  
 XX Human; colon cancer; cancer; tissue profiling; forensic; mapping;  
 XX Genetic analysis; diagnostic; antisense therapy; gene; ss.  
 XX Homo sapiens.  
 OS WO200229086-A2.  
 FN 11-APR-2002.  
 PD 02-OCT-2001; 2001WO-US030732.  
 PP 02-OCT-2000; 2000US-0237271P.  
 PR (FARB) BAYER CORP.  
 PI Burgess C, Astle JH, Carroll E, Catino TJ, Dwivedi P, Molino GA;  
 PI Thiagalingam A, Lewis MB;  
 PI WPI; 2002-426115/45.  
 DR New isolated nucleic acid that is differentially expressed in cancer  
 XX tissues useful for determining the presence of colon cancer in a cell or  
 PT tissue type, and in antisense therapy.  
 PS Claim 1; Fig 1; 796pp; English.  
 XX ABQ56306 to ABQ60787 represent isolated nucleic acids (I) differentially  
 CC expressed in cancer tissues. ABQ78993 to ABQ79004 represent proteins  
 CC encoded by the ABQ60776 to ABQ60787 nucleic acid sequences. (I) can be  
 CC used in antisense therapy. An antibody immunoreactive with a polypeptide  
 CC encoded by (I) is useful for detecting cancer in a patient sample, and  
 CC for detecting the presence or absence of a polynucleotide encoded by a  
 CC nucleic acid which hybridizes to (I) in a cell. A probe/primer derived  
 CC from (I) can be used for determining the presence of a nucleic acid which  
 CC hybridizes to (I), and for determining the phenotype of cells in a sample  
 CC of cells from a patient. (I) is useful for determining the presence of  
 CC colon cancer in a cell or tissue type, for determining the presence or  
 CC state of other type of cancer, in antisense therapy, to generate  
 CC macroarrays on a solid surface, to identify a chromosome on which the  
 CC corresponding gene resides, and in tissue profiling, forensics, genetic  
 CC analysis, mapping and diagnostic applications. (I) can be used to raise  
 CC antibodies, and to screen for peptide analogues and antagonists  
 XX Sequence 669 BP; 177 A; 169 C; 156 G; 167 T; 0 U; 0 Other;  
 SQ Query Match 30.9%; Score 635; DB 6; Length 669;  
 Best Local Similarity 99.6%; Pred. No. 6.1e-115;  
 Matches 668; Conservative 0; Mismatches 0; Indels 3; Gaps 3;





Db 760 AATAGAAAAACCCACGGGAACCAAGAGGACGGGCTTGCTTGAGCCTTACTCCAA 819  
 Qy 1892 GTCAGACCCAAAGGTCAATGCACAGAGACCCCGGGTGGTGAGCGCTGGCT 1943  
 Db 820 GT--AGCCCCCAAGGCTATGCACAGAAACCCCGGTAGGGAAACCTAGACT 869

Search completed: March 10, 2004, 11:00:13  
 Job time : 505.903 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 10, 2004, 09:19:42 ; Search time 93.7976 Seconds  
(without alignments)  
12152.440 Million cell updates/sec

Title: US-10-084-817-32

Perfect score: 2054  
Sequence: 1 cgggcctgcctgtccgc.....cattgaataaaaaaaaaa 2054

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA:\*\*

1: /cgn2\_6/ptodata/2/ina/5A.COMB.seq:\*\*

2: /cgn2\_6/ptodata/2/ina/5B.COMB.seq:\*\*

3: /cgn2\_6/ptodata/2/ina/6A.COMB.seq:\*\*

4: /cgn2\_6/ptodata/2/ina/6B.COMB.seq:\*\*

5: /cgn2\_6/ptodata/2/ina/PTCUS.COMB.seq:\*\*

6: /cgn2\_6/ptodata/2/ina/backfiles1.seq:\*\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2034	99.0	2036	3	US-08-923-454A-17
2	2010.8	97.9	2205	3	US-08-888-077A-41
3	493.4	24.0	495	4	US-09-702-705-897
4	493.4	24.0	495	4	US-09-736-457-897
5	493.4	24.0	495	4	US-09-614-124B-897
6	493.4	24.0	495	4	US-09-671-325-897
7	479.4	23.3	539	1	US-08-322-742-18
8	419	20.4	441	4	US-09-621-976-14193
9	381.8	18.6	1859	4	US-09-724-864-22
10	317	15.4	358	1	US-08-322-742-19
11	288	14.0	1835	3	US-08-923-454A-7
12	289	14.0	2040	4	US-09-075-460-4
13	288	14.0	2187	3	US-08-923-454A-23
14	288	14.0	2187	3	US-08-923-454A-24
15	288	14.0	2187	3	US-08-923-454A-30
16	288	14.0	2476	3	US-09-008-271A-23
17	224.8	10.9	1503	3	US-08-923-454A-5
18	224.8	10.9	2144	3	US-08-923-454A-28
19	212.4	10.3	1787	3	US-08-923-454A-3
20	212.4	10.3	2551	3	US-08-923-454A-26
21	119	5.8	732	3	US-08-923-454A-1
22	109.8	5.3	1980	1	US-08-350-741-1
23	109.8	5.3	1980	2	US-08-463-875A-1
24	102.4	5.0	1488	4	US-09-489-039A-5255
25	102.4	5.0	1488	4	US-09-489-039A-5193
26	101	4.9	558	4	US-08-976-259-53
27	97	4.7	1854	3	US-08-923-454A-13

28	92.2	4.5	111	1	US-08-322-742-4	Sequence 4, Appl1
29	89.8	4.4	1224	4	US-09-252-931A-8171	Sequence 8171, Ap
30	89.8	4.4	1380	4	US-08-252-931A-7981	Sequence 7981, Ap
31	89.8	4.4	1425	4	US-09-489-039A-6523	Sequence 6523, Ap
32	88.6	4.3	1368	4	US-09-711-164-245	Sequence 245, App
33	88	4.3	1436	4	US-09-199-637A-131	Sequence 131, App
34	87.6	4.3	1083	4	US-09-489-039A-6463	Sequence 6463, Ap
35	84.8	4.1	1428	4	US-09-252-931A-12923	Sequence 12923, A
36	84.8	4.1	1435	4	US-09-252-931A-12561	Sequence 12561, A
37	79.4	3.9	4403765	3	US-09-103-840A-2	Sequence 2, Appl1
38	79.4	3.9	4411529	3	US-09-103-840A-1	Sequence 1, Appl1
39	79.2	3.9	38675	4	US-08-311-731A-135	Sequence 135, App
40	75.6	3.7	891	4	US-08-252-931A-8283	Sequence 8283, Ap
41	71.4	3.5	1771	3	US-08-818-112-13	Sequence 13, Appl
42	71.4	3.5	1771	4	US-08-818-111-13	Sequence 13, Appl
43	71.4	3.5	1771	4	US-09-056-556-13	Sequence 13, Appl
44	71.4	3.5	1771	4	US-09-072-596-13	Sequence 13, Appl
45	71.4	3.5	1771	4	US-09-072-967-13	Sequence 13, Appl

ALIGNMENTS

RESULT 1  
US-08-923-454A-17  
; Sequence 17, Application US/08923454A  
; Patent No. 6004794  
; GENERAL INFORMATION:  
; APPLICANT: Creasy, Caretha  
; APPLICANT: Liivi, George  
; APPLICANT: Karkan, Eric  
; APPLICANT: Clinkbeard, Helen  
; APPLICANT: Browne, Michael  
; APPLICANT: Southan, Christopher  
; TITLE OF INVENTION: HUMAN SERINE PROTEASE  
; NUMBER OF SEQUENCES: 40  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Smithkline Beecham Corporation  
; STREET: 709 Swedeland Road  
; CITY: King of Prussia  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19406  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq Version 1.5  
; CURRENT APPLICATION DATA: US/08/923,454A  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION NUMBER: 60/025436  
; FILING DATE: 06-SEPT-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Baumeister, Kirk  
; REGISTRATION NUMBER: 33,833  
; REFERENCE/DOCKET NUMBER: P50547  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 610-270-5096  
; TELEFAX: 610-270-5090  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 17:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2036 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE:

[illegible]

## RESULT 2



```

Db 1456 CGCAGTATGTCAGGAGCTCAATTAAGGAGAAAGCCCTGGAACATGGTGTCCGCAG 1515
QY 1440 GGGTAATGAAGATATCATGATACAGTGTATCCCGAAGAAATGACCCATAGCAGAGGC 1499
Db 1516 GGGTAATGAAGATATCATGATACAGTGTATCCCGAAGAAATGACCCATAGCAGAGGC 1575
QY 1500 ATGAGCTGGACTTCATGTTTCCCTCAAGAGCTCTCCGCGGATGACGGATGAGGACTCTG 1559
Db 1576 ATGAGCTGGACTTCATGTTTCCCTCAAGAGCTCTCCGCGGATGACGGATGAGGACTCTG 1635
QY 1560 GGCTGCTGGAATAGACACTCAAGACTTTTGTGATGCTGCCATTTTGTGTTTGTGAGTGGAGACT 1619
Db 1636 GGCTGCTGGAATAGACACTCAAGACTTTTGTGATGCTGCCATTTTGTGTTTGTGAGTGGAGACT 1695
QY 1620 CCCTGGCCAAAGAGATCTCTCTTGTATGTTTGCAGGCAAAACAAATGTAATGTTGCAGAT 1679
Db 1696 CCCTGGCCAAAGAGATCTCTCTTGTATGTTTGCAGGCAAAACAAATGTAATGTTGCAGAT 1755
QY 1680 CCGCAGGAGAGCTCTGCCCTTCTGTATCTCTATGATGAGTGTGCTTTTCTTGGCCAG 1739
Db 1756 CCGCAGGAGAGCTCTGCCCTTCTGTATCTCTATGATGAGTGTGCTTTTCTTGGCCAG 1815
QY 1740 CTTCGGCCATTTCTGTTAGACAGTCAAGATTTTCTCTCTCTTAACTGAGTCAATC 1799
Db 1816 CTTCGGCCATTTCTGTTAGACAGTCAAGATTTTCTCTCTCTTAACTGAGTCAATC 1875
QY 1800 TTAGTCCAACTAATGACGTGATCAATGCGTAGATAGAAAGCCCGGAGGAGCCAGG 1859
Db 1876 TTAGTCCAACTAATGACGTGATCAATGCGTAGATAGAAAGCCCGGAGGAGCCAGG 1935
QY 1860 ATGGAGCTGGTGTGTTGTTGCTTTCTCAAGTACAGACCCCAAGGTCAATGCACAGAG 1919
Db 1936 ATGGAGCTGGTGTGTTGTTGCTTTCTCAAGTACAGACCCCAAGGTCAATGCACAGAG 1995
QY 1920 ACCCGGTGGTGTGAGCTGCTCTCTCAAGCGCGAAGTGTGCTTTTGAAGATCTC 1979
Db 1996 ACCCGGTGGTGTGAGCTGCTCTCTCAAGCGCGAAGTGTGCTTTTGAAGATNTN 2055
QY 1980 TTTGGAATGGGAGCAGATGACTCTGAGTTTGTAGCTATTAAGTACTCTTACACATG 2039
Db 2056 TTTGGAATGGGAGCAGATGACTCTGAGTTTGTAGCTATTAAGTACTCTTACACATG 2115
QY 2040 AAAAAAAAAAAAAA 2054
Db 2116 AAAAAAAAAAAAAA 2130

```

RESULT 3

```

US-09-702-705-897
; Sequence 897, Application US/09702705
; Patent No. 6504010
; GENERAL INFORMATION:
; APPLICANT: Wang, Tonglong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Lijun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C14
; CURRENT APPLICATION NUMBER: US/09/702,705
; CURRENT FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 1833
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 897
; LENGTH: 495
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-702-705-897

```

```

Query Match 24.0%; Score 493.4; DB 4; Length 495;
Best Local Similarity 99.8%; Pred. No. 5.8e-104;
Matches 494; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1094 CTGGAATCTCTTTTGAATCCCATCTGATAAGATTAAAAAGTTTCTTCCACGAGTCCCATG 1153
Db 1 CTGGAATCTCTTTTGAATCCCATCTGATAAGATTAAAAAGTTTCTTCCACGAGTCCCATG 60
QY 1154 ACCGACAGCCAAAGGAAGCCATCACAAGAGAGTATATTTGGTATCCGATGATGT 1213
Db 61 ACCGACAGCCAAAGGAAGCCATCACAAGAGAGTATATTTGGTATCCGATGATGT 120
QY 1214 CACTCACGTCCAGCAAAAGCCAAAGAGCTGAAGAGACCGGCACCGGGACTTCCACAGACGTGA 1273
Db 121 CACTCACGTCCAGCAAAAGCCAAAGAGCTGAAGAGACCGGCACCGGGACTTCCACAGACGTGA 180
QY 1274 TCTCAGGAGCGTATATAATTTGAAGTAAATTCCTGTATACCCCGAGAGCTGGTGGTCTCA 1333
Db 181 TCTCAGGAGCGTATATAATTTGAAGTAAATTCCTGTATACCCCGAGAGCTGGTGGTCTCA 240
QY 1334 AGAAAAACGAGCTCATAATCAGCATCAATGAGCAGTCCGTGGTCTCCGCCAATGATGTCA 1393
Db 241 AGAAAAACGAGCTCATAATCAGCATCAATGAGCAGTCCGTGGTCTCCGCCAATGATGTCA 300
QY 1394 CGCAGCTCATTAAGGGAAGACCCCTGAACATGCTGTGCTCCGACGGGTAAAGAGATA 1453
Db 301 CGCAGCTCATTAAGGGAAGACCCCTGAACATGCTGTGCTCCGACGGGTAAAGAGATA 360
QY 1454 TCATGATCACAGTGAATCCCGAAGAAATTTGACCCATAGCAGAGGATGAGTGGACTTC 1513
Db 361 TCATGATCACAGTGAATCCCGAAGAAATTTGACCCATAGCAGAGGATGAGTGGACTTC 420
QY 1514 ATGTTTCCCTCAAGACTCTCCGCGGATGAGGATGAGGACTCTGGGCTGCTGGAATAG 1573
Db 421 ATGTTTCCCTCAAGACTCTCCGCGGATGAGGATGAGGACTCTGGGCTGCTGGAATAG 480
QY 1574 GACACTCAAGACTTTT 1588
Db 481 GACACTCAAGACTTTT 495

```

RESULT 4

```

US-09-736-457-897
; Sequence 897, Application US/09736457
; Patent No. 6509448
; GENERAL INFORMATION:
; APPLICANT: Wang, Tonglong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Lijun
; APPLICANT: Wang, Aijun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C15
; CURRENT APPLICATION NUMBER: US/09/736,457
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 1864
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 897
; LENGTH: 495
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-736-457-897

```

```

Query Match 24.0%; Score 493.4; DB 4; Length 495;
Best Local Similarity 99.8%; Pred. No. 5.8e-104;
Matches 494; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```











QY 1141 ACGAGTCCATGACCGACAGGCCAAAGGAAAGCCATCACCAGAAAGATATATGGT 1200  
DB 1280 CGTGGGAAAAGAAATTCCTCTCCGGATCATGGTCCCGCGCTACATGGG 1339  
QY 1201 ATCCGATGATGATCATCAGTCCAGCAAGAAAGCTGAAAGACCGGACCGGAC 1260  
DB 1340 GTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1399  
QY 1261 TTTCCAGAGCTGATCTCAGGAGCGTATATATATATATATATATATATATAT 1320  
DB 1400 TTTCCGATGTTTCCAGATGTTTCTATCATCAATCAATCAATCAATCAATCA 1459  
QY 1321 GCTGGTGGTCTCAAGAAAGACGCTCATATCAATCAATCAATCAATCAATCA 1380  
DB 1460 CGGCTGGTCTGCGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1519  
QY 1381 GCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1440  
DB 1520 GCTGAAGATGTTTATGAAGTGTTCGAACCAATCCCAATCCCAATCCCAAT 1579  
QY 1441 GGTAAATGAAGATATCATGATCATGATGATGATGATGATGATGATGATGAT 1484  
DB 1580 GGACGAGAAACACTGACCTTATATGATGATGATGATGATGATGATGATGAT 1623

## RESULT 12

US-09-075-460-4

; Sequence 4, Application US/09075460A

; Patent No. 6489136

; GENERAL INFORMATION:

; APPLICANT: Zervos, Antonis S.

; TITLE OF INVENTION: CELL PROLIFERATION RELATED GENES

; FILE REFERENCE: 10284/004001

; CURRENT APPLICATION NUMBER: US/09/075,460A

; CURRENT FILING DATE: 1998-05-08

; EARLIER FILING DATE: 1997-05-09

; NUMBER OF SEQ ID NOS: 16

; SOFTWARE: FASTSEQ for Windows Version 3.0

; SEQ ID NO 4

; LENGTH: 2040

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (248) ... (1834)

US-09-075-460-4

Query Match 14.0%; Score 288; DB 4; Length 2040;

Best Local Similarity 57.3%; Pred. No. 1.1e-56;

Matches 541; Conservative 0; Mismatches 400; Indels 3; Gaps 1;

QY 544 CGCCATAAATATAACTTTATCGGAGCTGGTGGAGAGATCGCCCTCGCGTGGTTCAT 603  
DB 890 CGGAGTCAGTACAACTTCATCGAGATGTTGGTGGAGAGACAGACCTCGCGTGGTCTAT 949  
QY 604 ATCGAATGTTTCGACGCTTCGTTTCTAAAGAGAGGTCGCGGTGGTGGTGGTCT 663  
DB 950 ATCGAGATCTCGACCGGACCTTCTTGGCGCGAGGTCCTTATCTCGAACGGCTCA 1009  
QY 664 GGGTTTATGTTGTCGGAAGATGACACTGATCGTGACAAATGCCACGCTGGTGACCAACAG 723  
DB 1010 GGATTCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1069  
QY 724 CACCGGTCAAAGTTGAGCTGAAGAGCGTCCCTACGAGGCAAAATCAAGATGTTG 783  
DB 1070 CGCAGATCGGTGAGACTGCTAAGCGGCGACACGATATAGGCGCGTGGTCAAGCTGTG 1129  
QY 784 GATGAGAAAGACAGATCGCACTCATCAAAATGACCAACGAGGCAAGCTGCTGCTG 843  
DB 1130 GATCCCGTGCAGACATCGCAACGCTGAGGATTCAGACTAAGGACCTCTCCCAAGCTG 1189  
QY 844 CTGCTTGGCGCTCTCTCAGAGCTCGCGCGGAGAGTTCGTGGTGGTGGTGGTGG 903

DB 1190 CCTCTGGAGCTCAGCTGATGTCCGGCAAGGGAGTTTGTGTGCTATGGAGTCCC 1249  
QY 904 TTTTCCCTTCAAACACAGTCCACCGGGATCGGTAGGACCAACCCAGGAGCGCGCAAA 963  
DB 1250 TTTGCACTGAGAACACGATCACATCCGGCATTTGTAGCTCTGCTCAGCGTCCAGCCAGA 1309  
QY 964 GAGCTGGGGTCCGCAACTCAGACATGAGTACATCCAGACCGACGCCATCATCAACTAT 1023  
DB 1310 GACTGGGACTCCCGCAACCAATGTGGAATACATTCAAACTGATGACGCTATTTT 1369  
QY 1024 GGAACCTGGAGGCGCGTGTAGTAACTGAGCGGTGAAGTGAATTAACACTTTG 1083  
DB 1370 GGAACCTGAGAGTCCCTGGTTAACTGATGGGAGGTGATTTGAGAGTGAACACCATG 1429  
QY 1084 AAAGTGAAGTGAATCTCTTTTGGCAATCCATCTGATAAGATTAATAAGATTCCT---C 1140  
DB 1430 AAGTCAAGTGAATCTCTTTTGGCAATCCCTTCTGATCGTCTCGAGAGTTTCTGCAT 1489  
QY 1141 ACGAGTCCCATGACCGACAGGCCAAAGAAAGCCATCAACCAAGAAAGTATATTTGGT 1200  
DB 1490 CGTGGGAAAGAAAGAAATTCCTCTCCGGAATCAGTGGTCCCGCGCTACATTTGGG 1549  
QY 1201 ATCCGAATGATGTCACTCAGTCCAGCAAGCAAGAGCTGAAGACCGGACCGGAC 1260  
DB 1550 GTGATGATGCTGACCTGAGTCCAGCATCTTGTCTGACTACAGCTTCGAGAACCAAGC 1609  
QY 1261 TTCCAGACGTGATCTCAGGAGCGTATATAATTTGAAGTAAATTCCTGATACCCAGAGAA 1320  
DB 1610 TTTCCGATGTTTCCAGCATGTTGATGATGATGATGATGATGATGATGATGATGAT 1669  
QY 1321 GCTGGTGGTCTCAAGAAAGACGCTCATATATCAGCATCAATCAGCATCAATGAGTGGTCTCC 1380  
DB 1670 CGGCTGGTCTCGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1729  
QY 1381 GCAATGATGTCAGGAGCTGATTAAGAAAGGAAAGACCCCTGAAATGTTGGTGGTGGTGG 1440  
DB 1730 GCTGAAGATGTTTATGAAGCTGTTTGAACCAATCCAGTTCGAGTGGCAGTGCAGATCCGCGG 1789  
QY 1441 GGTAAATGAAGATCATGATCAGCATGATTCCTCGAAGAAATGA 1484  
DB 1790 GGACGAGAAACACTGACCTTATATGTCACCTCGAGTGCACAGA 1833

## RESULT 13

US-08-923-454A-23

; Sequence 23, Application US/08923454A

; Patent No. 6004794

; GENERAL INFORMATION:

; APPLICANT: Creasy, Caretha

; APPLICANT: Livi, George

; APPLICANT: Karkan, Eric

; APPLICANT: Clinkenbeard, Helen

; APPLICANT: Browne, Michael

; APPLICANT: Southan, Christopher

; TITLE OF INVENTION: HUMAN SERINE PROTEASE

; NUMBER OF SEQUENCES: 40

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Smithline Beecham Corporation

; STREET: 709 Swedeland Road

; CITY: King of Prussia

; STATE: PA

; COUNTRY: USA

; ZIP: 19406

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FASTSEQ Version 1.5

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/923,454A

; FILING DATE:

; CLASSIFICATION: 435

```
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/025436
; FILING DATE: 06-SEPT-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Baumeister, Kirk
; REGISTRATION NUMBER: 33,833
; REFERENCE/DOCKET NUMBER: P50547
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5096
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2187 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE:
; ORIGINAL SOURCE:
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 603...1976
; OTHER INFORMATION:
;
; US-08-923-454A-23

Query Match 14.0%; Score 288; DB 3; Length 2187;
Best Local Similarity 57.3%; Pred. No. 1.2e-56;
Matches 541; Conservative 0; Mismatches 400; Indels 3; Gaps 1;

QY 544 GCCATAATAAATTCATCGGAGCGTGTGGAGAGATCGCCCTGCGGTGTTTCAT 603
DB 1032 CGGAGTCAGTACAACTTCATCGAGATGTTGGAGAGACAGACCTGCGGTCTAT 1091
QY 604 ATCGAATGTTTCGAAGCTTCGTTTTCATACGAGAGTGGCGGTGAGTGGTCT 663
DB 1092 ATCGAGATCCTGGACCGGACCCCTTCTTGGCGCGAGTCCCTATCTCGAAGCGTCA 1151
QY 664 GGGTTTATTGTGCGAAGATGAGTATCGTGACAAATGCCACGCTGTGACCAACAAG 723
DB 1152 GGATTCGTGGTGGCTGCCGATGGCTCATTTGTCACCAACGCCCATGTGGTGTGATCGG 1211
QY 724 CACCGGTCAAGTTGAGCTGAAGACGTTGCCATTCAGAGCCAAATCAGAGATGTG 783
DB 1212 CGCAGATCCGTGTGAGATGCTAAGCGGCGACACGTATAGGCGGTGTCACAGCTGTG 1271
QY 784 GATGAGAAAGCAGACATCGCACTCATCAAAATTGACCACCGAGGCGCAAGCTGCTGTG 843
DB 1272 GATCCCGTGGCAGACATCGCAAGCTGAGATTTCAGACTAAGGAGCCTCTCCCAAGCTG 1331
QY 844 CTGCTTGGCGCTCCTCAGAGCTGCGCGGGAGAGTTCGTGTCGCCATCGGAGCCG 903
DB 1332 CCTCGGAGCTGAGCTGATGTCGGCAAGGGAGTTCGTGTCATGGGAGTCCC 1391
QY 904 TTTTCCCTTCAAAACACAGTCACCAACCGGATCGTGAGCACCAACCGAGCGGCAAA 963
DB 1392 TTTGACATGCAACACAGTACATCGGCGATTTAGCTCTGCTCAGCTCCAGCCAGA 1451
QY 964 GAGCTGGGGTCCGCAACTCAGACATGGAATCATCCAGACCGGACCCATCAACTAT 1023
DB 1452 GACCTGGGACTCCCCCAACCAATGTGGAAATCATTCAAACTGATGAGCTATGATTTT 1511
QY 1024 GGAAACTCGGGAGGCCGTTAGTAAACCTGGACGTTGAATGATTTGAATTAACACTTTG 1083
DB 1512 GGAAACTCTGGAGTCCCTGTTAACTGGATGGGAGGTGATTTGAGTGAACACCATG 1571
QY 1084 AAAGTGACGCTGGATTCCTTTTGAATCCCATCTGATAAGATTAAGATTCCT--C 1140
DB 1572 AAGTCAACGCTGGAAATCTCTTTTGGCATCCCTTCTGATCGTCTCGAGAGTTTCTGCAT 1631
QY 1141 ACGGAGTCCCATGACCGACAGCGCCAAAGGAAAGCCATCAACCAAGAAAGATATTGGT 1200

; RESULT 14
; US-08-923-454A-24
; Sequence 24, Application US/08923454A
; Patent No 6004794
; GENERAL INFORMATION:
; APPLICANT: Creasy, Caratha
; APPLICANT: Livi, George
; APPLICANT: Karran, Eric
; APPLICANT: Clinkenbeard, Helen
; APPLICANT: Browne, Michael
; APPLICANT: Southan, Christopher
; TITLE OF INVENTION: HUMAN SERINE PROTEASE
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/923,454A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/025436
; FILING DATE: 06-SEPT-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Baumeister, Kirk
; REGISTRATION NUMBER: 33,833
; REFERENCE/DOCKET NUMBER: P50547
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5096
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2187 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO

DB 1632 CGTGGGAAAGAAAGATTCCTCCTCCGGAATCAGTGGGTCCAGCGCGGTACATTGGG 1691
QY 1201 ATCCGAATGATGTCTACTCAGCTCCAGCAAGCCAAAGAGCTGAAGGACCGGCACCGGAC 1260
DB 1692 GTGATGATGCTGACCCCTGAGTCCAGCATCCTTTGCTGAATACAGCTTCGAGAACCAGC 1751
QY 1261 TTCCAGAGCGTATCTCAGGAGCGTATATAAATTGAAGTAAATTCCTGTATACCCAGCAGAA 1320
DB 1752 TTTCCCGATGTTTCAGCATGGTGTACTCATCCATAAAAGTCAATCCTGGGCTCCCTGCACAC 1811
QY 1321 GCTGCTGCTCTCAAGGAAACGACGTCAATATCAGCATCAATGACAGTCCGTGGTCTCC 1380
DB 1812 CGGCTGGTCTCGGCTGATGATGATTTTGGCCATTTGGGAGCAGATGTTACAAAAT 1871
QY 1381 GCCAATGATGTGAGCGACGTCTATTAAGGAGGAAAGCACCCCTGAACATCGTGGTCCGCGAGG 1440
DB 1872 GCTGAAGATGTTTATGAAGCTGTTGCAACCCAAATCCAGTTGGCAGTGCAGATCCGGCGG 1931
QY 1441 GGTAAATGAAGATATCATGATCAGATGATTCAGATTCCTCCGGAAGAAATTGA 1484
DB 1932 GGACGAGAAACACTGACCTTATATGTGACCCCTGAGTCAACA 1975
```

FRAGMENT TYPE:  
; ORIGINAL SOURCE:  
; FEATURE:  
; NAME/KEY: Coding Sequence  
; LOCATION: 603...1976  
; OTHER INFORMATION:  
US-08-923-454A-24

Query Match  
Best Local Similarity 14.0%; Score 288; DB 3; Length 2187;  
Matches 541; Conservative 0; Mismatches 400; Indels 3; Gaps 1;

QY 544 CGCCATAAATAAATTAACCTTATCGCGGACGCTGGTGAGAGATCGCCCTCGCGTGGTTTCAT 603  
Db 1032 CGGAGTCAGTACAACTTCACTCGCAGATGTGGTGAGAGACAGCACCTCGCGTGGTCTAT 1091  
QY 604 ATCGAATTGTTTCGCAAGCTTCGGTTTTCTAAACGAGAGGTGCGCGTGGCTAGTGGGTCT 663  
Db 1092 ATCGAGATCTCGACCGCACCTTTCTGGCGCGGAGGTCCCTATCTCGAACGGCTCA 1151  
QY 664 GGGTTTATTTGTTCGAGAGTGAAGTGAATGCTGACAAATGCCACGGTGGTGAACCAAG 723  
Db 1152 GGATTCGTGGTGGCTCGCGATGGCTCAATGTCAACACGCCCATGTGGTGGCTGATCGG 1211  
QY 724 CACCGGGTCAAAAGTTGAGCTGAAGAACGGTGCCACTTACGAAGCCAAATCAAGGATGTG 783  
Db 1212 CGCAGAGTCGGTGGAGCTGCTAAGCGCGACACGATATGAGCGCGTGGTTCACAGCTGTG 1271  
QY 784 GATGAGAAAGCAGACATCGCACTCACTCAAAATTGACACAGGCAAGCTGCTGCTGTG 843  
Db 1272 GATCCGGTGCAGACATCGCAACGCTGAGGATTCAGACTAAGGAGCCCTCTCCCAACGGTG 1331  
QY 844 CTGCTTGGCGCTCCTCAGAGCTGCGGCGCGGAGATTGCTGGTCCCATCGGAAGCCCG 903  
Db 1332 CCGTGGGAGCTCAGCTGATGCTCCGCAAGGGAGTTGTTGTGCCATCGGAAGTCCC 1391  
QY 904 TTTTCCCTTCAAAACAGTCAACACCGGATCGTGACCAACCCAGCGAGGCGGCAAA 963  
Db 1392 TTGCACTGAGAACACAGTACATCCGGCATTTGTTAGCTGTCTGACGGTCCAGCCAGA 1451  
QY 964 GAGCTGGGGCTCGCAACTCAGACATGAGCTATACATCCAGACCGGCGCCATCATCAACTAT 1023  
Db 1452 GACCTGGGACTCCCAACCAATGTGAATACATTCAAACTGATGACGCTATTGATTTT 1511  
QY 1024 GGAAGTGGGAGCGGTTAGTAAACCTCGACGCTGAAGTGAATTAACACTTTG 1083  
Db 1512 GGAAGTCTGGAGTCCCTGGTTAACTGATGGGAGGTTGATGGAGTGAACACCATG 1571  
QY 1084 AAAGTCAAGCTGGAATCTCCTTTGCAATCCATCTGATAGATTAAAGTTCTCT--C 1140  
Db 1572 AAGGTCACAGCTGGAATCTCCTTTGCCATCCTCTGATGCTCTGAGAGTTTCTGCAT 1631  
QY 1141 ACGGAGTCCATGACGACAGGCAAGGCAAGAAAGCCATCACCAGAGAAATATATTCGT 1200  
Db 1632 CGTGGGAAAGAAAGAAATTCCTCTCGGAATCAGTGGTCCGAGCGGCTACATTTGG 1691  
QY 1201 ATCCGAATGATGTCACTCAGTCCAGCAAAAGCCAAAGAGCTGAAGGACCGGACCGGAC 1260  
Db 1692 GTGATGATGTCAGCCCTGAGTCCAGCATCTTGTGAACCTACAGCTTCGAGAACCAAGC 1751  
QY 1261 TTCCAGAGCTGATCTCAGAGCGGTATATAATTAAGTAATTCCTGATACCCACGAGNA 1320  
Db 1752 TTTCCGAGTTCAGATGTTGACTCATCATTAAGTCACTCTGGGCTCCCTCGACAC 1811  
QY 1321 GCTGTGGTCTCAAGGAAACGAGCTCATATATGACATCAATGAGACGTCGGTGGTCTCC 1380  
Db 1812 CGGGCTGGTCTCGCGGCTGGTGTGATTTGGCCATTTGGGAGCAGATGATGATACAAAT 1871  
QY 1381 GCCAATGATGTGAGGAGCTCATTAAGAGGAGAGACCCCTGAACATGCTGTCGCGAGG 1440  
Db 1872 GCTGAAGATGTTTATGAAGCTGTTCGAACCCCAATCCCAAGTGGGAGTGAAGATCCGCG 1931  
QY 1441 GGTAAATGAAGATATCATGATCAGAGTATTCGGAAGAAATTTGA 1484

Db 1932 GGACGAGAAACACTGACCTTATATGTGACCCCTGAGGTACAGAGA 1975

RESULT 15  
US-08-923-454A-30  
; Sequence 30, Application US/08923454A  
; Patent No. 6004794  
; GENERAL INFORMATION:  
; APPLICANT: Creasy, Caretha  
; APPLICANT: Livi, George  
; APPLICANT: Karian, Eric  
; APPLICANT: Clinkenbeard, Helen  
; APPLICANT: Browne, Michael  
; APPLICANT: Southan, Christopher  
; TITLE OF INVENTION: HUMAN SERINE PROTEASE  
; NUMBER OF SEQUENCES: 40  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SmithKline Beecham Corporation  
; STREET: 709 Swedeland Road  
; CITY: King of Prussia  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19406  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq Version 1.5  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/923,454A  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/025436  
; FILING DATE: 06-SEPT-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Baumeister, Kirk  
; REGISTRATION NUMBER: 33,833  
; REFERENCE/DOCKET NUMBER: P50547  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 610-270-5096  
; TELEFAX: 610-270-5090  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 30:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2187 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE:  
; ORIGINAL SOURCE:  
; FEATURE: Polymorphic variants at 672 and 1435  
; FEATURE: aa24=Arg/Cys aa278=Ala/Val  
; NAME/KEY: Coding Sequence  
; LOCATION: 603...1976  
; OTHER INFORMATION:  
US-08-923-454A-30

Query Match  
Best Local Similarity 14.0%; Score 288; DB 3; Length 2187;  
Matches 541; Conservative 0; Mismatches 400; Indels 3; Gaps 1;

QY 544 CGCCATAAATAAATTAACCTTATCGCGGACGCTGGTGAGAGATCGCCCTCGCGTGGTTTCAT 603  
Db 1032 CGGAGTCAGTACAACTTCACTCGCAGATGTGGTGAGAGACAGCACCTCGCGTGGTCTAT 1091  
QY 604 ATCGAATTGTTTCGCAAGCTTCGGTTTTCTAAACGAGAGGTGCGCGTGGCTAGTGGGTCT 663  
Db 1092 ATCGAGATCTCGACCGCACCTTTCTGGCGCGGAGGTCCCTATCTCGAACGGCTCA 1151



QY	664	GGGTTTATTGTGCGAAGATCGACTGATCTGACAAATGCCACGCTGGTGACCAACAAG	723
Db	1152	GGATTCTGGTGGCTGCCGATGGGCTCATGTGACCAACGCCCATGTGGTGGCTGATCGG	1211
QY	724	CACCGGFTCAAGTTGAGCTGAAGAACGGTGGCCACTTACGAAGCCAAATCAAGGATGTG	783
Db	1212	CGCAGAGTCCGTGTGAGACTGTAAAGCGCGACACGATATGAGCCGTGGTACAGCTGTG	1271
QY	784	GATGAGAAAGCAGACATCGCACTCATCAAAATTTGACCACAGGGCAAGCTGCCTGTCTCG	843
Db	1272	GATCCCGTGGCAGACATGCAACGCTGAGGATTCAGACTAAGAGCCCTCTCCCAACGCTG	1331
QY	844	CTGCTTGGCCCGCTCTCAGAGCTGCGCGCGGAGAGTTGCTGGTCCGCATCGGAAGCCCG	903
Db	1332	CCTCTGGACCTCAGCTGATGTCGGGCAAGGGAGTTTGTGTTGCCATGGGAAGTCCC	1391
QY	904	TTTTCCCTTCAAAACACAGTCAACCCGGGATCGTGAGCACCACCCAGCGAGGCGGCAAA	963
Db	1392	TTTGCACTGCGAGAACACGATCACATCCGGCATTGTTAGCTCTCTCAGGTCAGCCAGA	1451
QY	964	GAGCTGGGCTCCGCAACTCAGACATGGACTACATCCAGACCGACGCCATCATCACTAT	1023
Db	1452	GACCTGGGACTCCCCCAACCAATGTGGAATACATTCAAACTGATGAGCTATTGATTTT	1511
QY	1024	GGAAACTCGGAGGCGCGTTAGTAAACCTGGACGGTGAAGTGAATGGAAATTAACACTTTG	1083
Db	1512	GGAAACTCTGGAGTCCCTGGTTAACTTGAATGGGAGGTGATTGGAGTGAACACCATG	1571
QY	1084	AAAGTGACAGCTGGAATCTCTTTCCATCCCATCTGATAAGATTAAAGTTTCCT---C	1140
Db	1572	AAAGTGACAGCTGGAATCTCTTTGGCATCCCTTTGATCGCTTCGAGAGTTCTGCTAT	1631
QY	1141	ACGAGTCCCATGACCGACAGGCCAAAGAAAGCCATCACCAGAAAGATATATTGGT	1200
Db	1632	CGTGGGAAAGAGAGATTCTCTCCGGAATCAGTGGGTCCAGCGGCGCTACATTGGG	1691
QY	1201	ATCCGAATGATGTCACTCAGCTCCAGCAAGCAAGAGCTGAAGACCGGCACCGGAC	1260
Db	1692	GTGATGATGCTGACCTGAGTCCAGCATCTTGTGTAATACAGCTTCGAGAACCAAGC	1751
QY	1261	TTCCAGACGATGATCTCAGGAGCGTATATAATTGAAGTAAATTCCTGATATCCCGAGAA	1320
Db	1752	TTTCCGATGTTACGATGGTGTACTCATCCATAAAGTCATCTCTGGGCTCCCTGCACAC	1811
QY	1321	GCTGGTGTCTCAGGAAACACGCTCATATAATCAGCATCAATGACAGTCCGTTGCTCC	1380
Db	1812	CGGCTGCTCTCGGCTGGTGTGATGATTTTGGCCATTGGGAGCAGATGGTACAAAT	1871
QY	1381	GCCATGATGTCAGCAGCTCATTTAAAGGAAAGCACCTGAAATGGTGGTCCGAGG	1440
Db	1872	GCTGAAGATGTTTATGAAGCTGTCGAAACCAATCCAGTTGGCAGTGCAGATCCGCGG	1931
QY	1441	GGTAATGAAGATATCATGATCACAGTGAATCCCGAAGAAATGA	1484
Db	1932	GGACGAGAAACACTGACCTTATATGTGACCCCTGAGGTCACAGA	1975

Search completed: March 11, 2004, 14:41:55  
Job time : 100.798 secs







Db	1922	TGGCAGCTGTCGTCTTTTCTGCTTTTCCTCAAGTCAGCACCCAAAAGTCAATGACACAGAGA	1981
QY	1921	CCCCGGGTGGGTGAGCGCTGGCTTCTCAACCGGCCGAGTGGCCCTTTTATGGAATCTCT	1980
Db	1982	CCCGGGGTGGGTGAGCGCTGGCTTCTCAACCGGCCGAGTGGCCCTTTTATGGAATCTCT	2041
QY	1981	TTGGAATTGGGAGGACGATGACTGATGATTGAGCTATTAAGTACTTCAACACTTGA	2040
Db	2042	TTGGAATTGGGAGGACGATGACTGATGATTGAGCTATTAAGTACTTCAACACTTGA	2101
QY	2041	AAAAAAAAAAAAA	2054
Db	2102	AAAAAAAAAAAAA	2115

```

RESULT 3
US-10-264-049-135
: Sequence 135, Application US/10264049
: Publication No. US20040005579a1
: GENERAL INFORMATION:
:   APPLICANT: Birse et al.
:   TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
:   FILE REFERENCE: PA133p1
:   CURRENT APPLICATION NUMBER: US/10/264,049
:   CURRENT FILING DATE: 2002-10-04
:   PRIOR APPLICATION NUMBER: PCT/US01/18569
:   PRIOR FILING DATE: 2001-06-07
:   PRIOR APPLICATION NUMBER: US 60/209,467
:   PRIOR FILING DATE: 2000-06-07
:   NUMBER OF SEQ ID NOS: 4360
:   SOFTWARE: PatentIn Ver. 3.1
:   SEQ ID NO 135
:   LENGTH: 2157
:   TYPE: DNA
:   ORGANISM: Homo sapiens
:   FEATURE:
:     NAME/KEY: misc feature
:     LOCATION: (1505)..(1505)
:     OTHER INFORMATION: n equals a,t,g, or c
:     FEATURES:
:       NAME/KEY: misc feature
:       LOCATION: (2147)..(2147)
:     OTHER INFORMATION: n equals a,t,g, or c
: US-10-264-049-135

```

Query Match	99.2%	Score 2037.4	DB 15	Length 2157
Best Local Similarity	99.8%	Pred. No. 0		
Matches 2049	Conservative	1	Mismatches 3	Indels 1
				Gaps 1

[illegible]

QY	361	GGCGAAGCGGCGCCTCTGTGTGTGTCGGCAGACGAGCGAGCGGTGTGGCGACGCAACGCAAC	420
Db	425	GGCGAAGCGGCGCCTCTGTGTGTGCGCA - CAGCGAGCGGTGTGTGGCGACGCAACGCAAC	483
QY	421	ACCTTACGCCAACCTTGTGCGCAGCTGTGGGCGCGCCAGCGCGCCTCCGAGAGGCTGTGCACCGG	480
Db	484	ACCTTACGCCAACCTTGTGCGCAGCTGTGGGCGCGCCAGCGCGCCTCCGAGAGGCTGTGCACCGG	543
QY	481	CCGCCGCTCATCGTCCCTGCGACCGCGAGCCTGCCGCCAAGGGCAGAAAGATCCCAACAGT	540
Db	544	CCGCCGCTCATCGTCCCTGCGACCGCGAGCCTGCCGCCAAGGGCAGAAAGATCCCAACAGT	603
QY	541	TTGGCCCAATAATATTAATTATTCGGCGGAGAGTGGTGGAGAAAGATGCCCTCCCGGT	600
Db	604	TTGGCCCAATAATATTAATTATTCGGCGGAGAGTGGTGGAGAAAGATGCCCTCCCGGT	663
QY	601	CATATCGAATTGTTTGGCAAGCTTCCGTTTCTTAAACGAGAGTGC CGGTGCTAGTGGG	660
Db	664	CATATCGAATTGTTTGGCAAGCTTCCGTTTCTTAAACGAGAGTGC CGGTGCTAGTGGG	723
QY	661	TCTGGGTTTATTTGTGTGTGGAAAGATGTGATGTGTGAACAATGCCACGTGTGACCAAC	720
Db	724	TCTGGGTTTATTTGTGTGTGGAAAGATGTGATGTGTGAACAATGCCACGTGTGACCAAC	783
QY	721	AAGCACCGGGTCAAAAGTTGAGCTGAAGAACGGTGCCTTACGAAGCCAAAATCAAGAT	780
Db	784	AAGCACCGGGTCAAAAGTTGAGCTGAAGAACGGTGCCTTACGAAGCCAAAATCAAGAT	843
QY	781	GTGGATGAGAAAGCAGACATCCGACATCATATAAATTGACACACAGGGGCAAGCTGCCTGC	840
Db	844	GTGGATGAGAAAGCAGACATCCGACATCATATAAATTGACACACAGGGGCAAGCTGCCTGC	903
QY	841	CTGCTGCTTGGCCGCTCTCTCAAGCTGGGCGCGGAGAGTTCTGTGTGCATCCGAAAC	900
Db	904	CTGCTGCTTGGCCGCTCTCTCAAGCTGGGCGCGGAGAGTTCTGTGTGCATCCGAAAC	963
QY	901	CCGTTTTCCCTTCAAAACAAGTACCAACCGGGATGTGAGACACACACCGAGGAGGCGGC	960
Db	964	CCGTTTTCCCTTCAAAACAAGTACCAACCGGGATGTGAGACACACACCGAGGAGGCGGC	1023
QY	961	AAAGAGCTGGGCTCCGCAACTGACATGAGATGACATTCAGACCGAGCGCATCATCAAC	1020
Db	1024	AAAGAGCTGGGCTCCGCAACTGACATGAGATGACATTCAGACCGAGCGCATCATCAAC	1083
QY	1021	TATGAAACTCGGGAGGCGCGGTAGTAAACCTGAGCGGTGAAGTATGGAAATTAACT	1080
Db	1084	TATGAAACTCGGGAGGCGCGGTAGTAAACCTGAGCGGTGAAGTATGGAAATTAACT	1143
QY	1081	TTGAAAGTGACAGCTGGAAATCTCTTTGCAATCCCATCTGATTAAGATTAAAGTTCTTC	1140
Db	1144	TTGAAAGTGACAGCTGGAAATCTCTTTGCAATCCCATCTGATTAAGATTAAAGTTCTTC	1203
QY	1141	ACGAGTCCCATGACCGGACAGGCGCAAGGAAAGAAAGCCATCAACAAAGAAATATTTGT	1200
Db	1204	ACGAGTCCCATGACCGGACAGGCGCAAGGAAAGAAAGCCATCAACAAAGAAATATTTGT	1263
QY	1201	ATCCGAATGATGTCACTACGTCCAGCAAAACCAAGAGCTGAAGGACCCGCGAC	1260
Db	1264	ATCCGAATGATGTCACTACGTCCAGCAAAACCAAGAGCTGAAGGACCCGCGAC	1323
QY	1261	TTCCGAGAGGTGATCTCAGGAGGGTATATATTAAGTAATTCCTCGATACCCCGACAGAA	1320
Db	1324	TTCCGAGAGGTGATCTCAGGAGGGTATATATTAAGTAATTCCTCGATACCCCGACAGAA	1383
QY	1321	GCTGTGTGTCTCAGGAGAAACGACGTCAATATCAGATCAATGCAAGTCCGTGTCTCC	1380
Db	1384	GCTGTGTGTCTCAGGAGAAACGACGTCAATATCAGATCAATGCAAGTCCGTGTCTCC	1443
QY	1381	GCCATGTATGTCAAGCAGACGTCTTAAAGGAAAGACCCCTGAACATGGTGTCCAGG	1440
Db	1444	GCCATGTATGTCAAGCAGACGTCTTAAAGGAAAGACCCCTGAACATGGTGTCCAGG	1503
QY	1441	GGTAATGAAATATCATGATCAGATGATTCGAGAGAAATTGACCCCATAGGACAGGCA	1500

Db	1504	GNTATATAGATTCATGATGATACAGCTGATTTCCCGAGAAATTAAGCCATAGGACAGAGCA	1563
QY	1501	TGAGCTGACTTCATGTTTCCCTCAAAAGACTCTCCCTGATGACGATGAGACTCTGG	1560
Db	1564	TGAGCTGAGACTTCATGTTTCCCTCAAAAGACTCTCCCTGATGACGATGAGACTCTGG	1623
QY	1561	GCTGCTGGAAATAGAGACACTCAAGCTTTTGACTGCAATTTGTTGTTGTTCACTGAGACTC	1620
Db	1624	GCTGCTGGAAATAGAGACACTCAAGCTTTTGACTGCAATTTGTTGTTGTTCACTGAGACTC	1683
QY	1621	CCTGGCCACAGAACTCTTGTGATGATTGACGACAAACAAATGTAATGTTGACATC	1680
Db	1684	CCTGGCCACAGAACTCTTGTGATGATTGACGACAAACAAATGTAATGTTGACATC	1743
QY	1681	CGACAGGACAGAAAGCTGCGCCCTCGATACCTATGATATGACATGACTTTTCTTGCCAGC	1740
Db	1744	CGACAGGACAGAAAGCTGCGCCCTCTGATACCTATGATATGACATGACTTTTCTTGCCAGC	1803
QY	1741	TTGGGCCATCTTCTTAGACAGTCAAGATTTGTCTCTCTTTAATGATCATCT	1800
Db	1804	TTGGGCCATCTTCTTAGACAGTCAAGATTTGTCTCTCTTTAATGATCATCT	1863
QY	1801	TAGTCCAACTATATGACAGTGCATTAATACCGATAGTAAGAAAGCCCAACGGGAGCCACGA	1860
Db	1864	TAGTCCAACTATATGACAGTGCATTAATACCGATAGTAAGAAAGCCCAACGGGAGCCACGA	1923
QY	1861	TGGGACTGGTGTGTTGTGTCTTTTCTTCAAGTACGACCCCAAGGTCAATGACAGAGA	1920
Db	1924	TGGGACTGGTGTGTTGTGTCTTTTCTTCAAGTACGACCCCAAGGTCAATGACAGAGA	1983
QY	1921	CCCCGGGTGGGTGAGCGCTGAGCTCTCAAGAGGCGGAAAGTGGCCCTTTTAGAATCTCT	1980
Db	1984	CCCCGGGTGGGTGAGCGCTGAGCTCTCAAGAGGCGGAAAGTGGCCCTTTTAGAATCTCT	2043
QY	1981	TTGGAATTTGGAGACAGATGAATCTGATTTGAGCTATTTAAAGTACTTCTTACATTTGA	2040
Db	2044	TTGGAATTTGGAGACAGATGAATCTGATTTGAGCTATTTAAAGTACTTCTTACATTTGA	2103
QY	2041	AAAAAAAAAAAAA 2054	
Db	2104	AAAAAAAAAAAAA 2117	

```

RESULT 4
US-09-960-706-516
Sequence 516, Application US/09960706
Publication No. US20030134280A1
GENERAL INFORMATION:
APPLICANT: Mungar, William E.
TITLE OF INVENTION: Identifying Drugs for and Diagnosis of Benign Prostatic Hyperplasia
FILE REFERENCE: 44921-5029-01US
CURRENT APPLICATION NUMBER: US/09/960,706
CURRENT FILING DATE: 2001-09-24
PRIOR APPLICATION NUMBER: 60/223,323
PRIOR FILING DATE: 2000-08-07
PRIOR APPLICATION NUMBER: 09/873,319
PRIOR FILING DATE: 2001-06-05
NUMBER OF SEQ ID NOS: 1124
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 516
LENGTH: 2036
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: Genbank Accession No. US20030134280A1 D87258
US-09-960-706-516

```

Query Match	99.0%	Score 2034.4	DB 10	Length 2036
Best Similarity	100.0%	Pred. NO. 0		
Best Local	0	Mismatches	1	Indels 0
Matches 2035		Conservative		Gaps 0

[illegible]





QY 841 CTGCTGCTTGGCGGCTCCTCAGAGCTCGGGCGGGAGAGTTCTGGTGGCCATCGGAAGC 900  
 Db 841 CTGCTGCTTGGCGGCTCCTCAGAGCTCGGGCGGGAGAGTTCTGGTGGCCATCGGAAGC 900  
 QY 901 CGTTTTCCTTCAAAACACAGTACACACCGGGATCGTGACACACACCGAGGCGGC 960  
 Db 901 CGTTTTCCTTCAAAACACAGTACACACCGGGATCGTGACACACACCGAGGCGGC 960  
 QY 961 AAAGAGCTGGGGTCCGCAATCAGACATGAGCTACATCCAGACGCGCATCATCAAC 1020  
 Db 961 AAAGAGCTGGGGTCCGCAATCAGACATGAGCTACATCCAGACGCGCATCATCAAC 1020  
 QY 1021 TATGGAACCTCGGAGGCGCGTGTAGTAACTGACACCGGTGAAGTGAATTAACACT 1080  
 Db 1021 TATGGAACCTCGGAGGCGCGTGTAGTAACTGACACCGGTGAAGTGAATTAACACT 1080  
 QY 1081 TTGAAGTGAACCTGGAATCTCTTTGCAATCCATCTGTATGATTAAGAGTTCTTC 1140  
 Db 1081 TTGAAGTGAACCTGGAATCTCTTTGCAATCCATCTGTATGATTAAGAGTTCTTC 1140  
 QY 1141 ACGAGTCCCATCAGCAGGCGCAAGGAAAGCCATCACCAAGAAAGTATATGGT 1200  
 Db 1141 ACGAGTCCCATCAGCAGGCGCAAGGAAAGCCATCACCAAGAAAGTATATGGT 1200  
 QY 1201 ATCCGAATGATGTCACCTCAGCTCCAGCAAGCCAAAGAGCTGAAGACCGGACCGGAC 1260  
 Db 1201 ATCCGAATGATGTCACCTCAGCTCCAGCAAGCCAAAGAGCTGAAGACCGGACCGGAC 1260  
 QY 1261 TTCCCAAGCTGATCTCAGAGCGGTATATATTAAGTAATCTCTGATACCCAGAGAA 1320  
 Db 1261 TTCCCAAGCTGATCTCAGAGCGGTATATATTAAGTAATCTCTGATACCCAGAGAA 1320  
 QY 1321 GCTGGTGGTCTCAAGGAAACGACGTATATATCAGCATCAATGACATCAATGGACGCTGCC 1380  
 Db 1321 GCTGGTGGTCTCAAGGAAACGACGTATATATCAGCATCAATGACATCAATGGACGCTGCC 1380  
 QY 1381 GCCAATGATGTCAGCGACGTATTAAGGAAAGACCTGACATGATGTCGCGCAGG 1440  
 Db 1381 GCCAATGATGTCAGCGACGTATTAAGGAAAGACCTGACATGATGTCGCGCAGG 1440  
 QY 1441 GTPAATGAAGATATCATGATCAGAGTATCCCGAAGAAATGACCCATAGCAGGCA 1500  
 Db 1441 GTPAATGAAGATATCATGATCAGAGTATCCCGAAGAAATGACCCATAGCAGGCA 1500  
 QY 1501 TGAGCTGGACTTCATGTTCCCTCAAGACTCTCCGTTGATGACGGATGAGGACTCTGG 1560  
 Db 1501 TGAGCTGGACTTCATGTTCCCTCAAGACTCTCCGTTGATGACGGATGAGGACTCTGG 1560  
 QY 1561 GCTGCTGGAATAGACACTCAAGACTTTTGAATGACATTTTGTTCAGTGGAGACTC 1620  
 Db 1561 GCTGCTGGAATAGACACTCAAGACTTTTGAATGACATTTTGTTCAGTGGAGACTC 1620  
 QY 1621 CTTGGCCAAACAGAACTCTTCTGATGTTTGCAGGCAAAACAAATGTAATGTCAGATC 1680  
 Db 1621 CTTGGCCAAACAGAACTCTTCTGATGTTTGCAGGCAAAACAAATGTAATGTCAGATC 1680  
 QY 1681 CGCAGGAGAAAGCTCTGCCCTCTGTATCTATGATGATGATGATGATGATGATGATGATG 1740  
 Db 1681 CGCAGGAGAAAGCTCTGCCCTCTGTATCTATGATGATGATGATGATGATGATGATGATG 1740  
 QY 1741 TTGGGCCATTTCTTCTTAGACAGTCAAGTATGCTCTCTCTTAACTGAGTCATCATCT 1800  
 Db 1741 TTGGGCCATTTCTTCTTAGACAGTCAAGTATGCTCTCTCTTAACTGAGTCATCATCT 1800  
 QY 1801 TAGTCCAACTAATGCACTGATGATCAATGATGATGATGATGATGATGATGATGATGATG 1860  
 Db 1801 TAGTCCAACTAATGCACTGATGATCAATGATGATGATGATGATGATGATGATGATGATG 1860  
 QY 1861 TGGGACTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1920  
 Db 1861 TGGGACTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1920

QY 1921 CCCCGGTGGTGGAGCGCTGCTTCTCAAAACGCGCAAGTTGCTCTTTTAGGAATCTCT 1980  
 Db 1921 CCCCGGTGGTGGAGCGCTGCTTCTCAAAACGCGCAAGTTGCTCTTTTAGGAATCTCT 1980  
 QY 1981 TTGGAATTGGAGCAGCATGACTCTGAGTTTGGAGCTATTAAGTACTTCTTACA 2034  
 Db 1981 TTGGAATTGGAGCAGCATGACTCTGAGTTTGGAGCTATTAAGTACTTCTTACA 2034

RESULT 6  
 US-10-102-806-134  
 ; Sequence 134, Application US/10102806  
 ; Publication No. US20030054421A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Rosen et al.  
 ; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
 ; FILE REFERENCE: P103P1C1  
 ; CURRENT APPLICATION NUMBER: US/10/102,806  
 ; CURRENT FILING DATE: 2002-03-22  
 ; PRIOR APPLICATION NUMBER: 09/925,298  
 ; PRIOR FILING DATE: 2001-08-10  
 ; PRIOR APPLICATION NUMBER: PCT/US00/05881  
 ; PRIOR FILING DATE: 2000-03-08  
 ; PRIOR APPLICATION NUMBER: 60/124,270  
 ; PRIOR FILING DATE: 1999-03-12  
 ; NUMBER OF SEQ ID NOS: 846  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 134  
 ; LENGTH: 1855  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: misc\_feature  
 ; LOCATION: (1818)  
 ; OTHER INFORMATION: n equals a,t,g, or c  
 ; NAME/KEY: misc\_feature  
 ; LOCATION: (1845)  
 ; OTHER INFORMATION: n equals a,t,g, or c  
 ; US-10-102-806-134

Query Match 86.7%; Score 1780; DB 14; Length 1855;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 1802; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

QY 251 GCGGCGCGCGCGAGGCGCGCTGCGGCTGCGAGGCGCGCTGCGGCGGCGGCGGCTGC 310  
 Db 14 GCGGCGCGCGCGAGGCGCGCTGCGGCTGCGAGGCGCGCTGCGGCGGCGGCGGCTGC 73  
 QY 311 AGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 370  
 Db 74 AGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 133  
 QY 371 GCCTCTGT 430  
 Db 134 GCCTCTGT 192  
 QY 431 ACCTGTGCCAGCTGCG 490  
 Db 193 ACCTGTGCCAGCTGCG 252  
 QY 491 TCGTCTGT 550  
 Db 253 TCGTCTGT 312  
 QY 551 AATATAACTTTATCGCGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 610  
 Db 313 AATATAACTTTATCGCGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 372  
 QY 611 TGTTCGCAAGCTTCCGTTTCTTAAACGAGAGGTGCCGTTGCTAGTGGTCTTGGGTTTA 670  
 Db 373 TGTTCGCAAGCTTCCGTTTCTTAAACGAGAGGTGCCGTTGCTAGTGGTCTTGGGTTTA 432  
 QY 671 TTGTGTGCGAAGATGGAAGT 730

Db 433 TTGTGTCGGAAGATGAGATGATGTCGTAACAAATGCCCACTGGTGTGACACACAGACCGGG 492  
 Qy 731 TCAAAAGTTGAGCTGAAGAACCGTGCACCTTACGAAGCCAAAATCAAGGATGTGGATGAGA 790  
 Db 493 TCAAAAGTTGAGCTGAAGAACCGTGCACCTTACGAAGCCAAAATCAAGGATGTGGATGAGA 552  
 Qy 791 AAGCAGACATCGCATCTCATCAAAATTTGACACACAGGCGCAAGCTGCTCTGCTGCTTG 850  
 Db 553 AAGCAGACATCGCATCTCATCAAAATTTGACACACAGGCGCAAGCTGCTCTGCTGCTTG 612  
 Qy 851 GCGCTCTCAGAGCTGGGCGGCGAGAGTGTGCTGCGCATCGAAGCCGTTTCC 910  
 Db 613 GCGCTCTCAGAGCTGGGCGGCGAGAGTGTGCTGCGCATCGAAGCCGTTTCC 672  
 Qy 911 TTCAAAACACAGTACACCGGATCGTGAGACCAACCGAGCGGCGGCAAGAGCTGG 970  
 Db 673 TTCAAAACACAGTACACCGGATCGTGAGACCAACCGAGCGGCGGCAAGAGCTGG 732  
 Qy 971 GGCTCCGCAACTCAGACATGGACTATCATCAGACCGACGCATCATCAACTATGGAAC 1030  
 Db 733 GGCTCCGCAACTCAGACATGGACTATCATCAGACCGACGCATCATCAACTATGGAAC 792  
 Qy 1031 CGGAGGCGCGTTAGTAACCTGGACGTTGAAGTGAATGAACCTTTGAAGTGA 1090  
 Db 793 CGGAGGCGCGTTAGTAACCTGGACGTTGAAGTGAATGAACCTTTGAAGTGA 852  
 Qy 1091 CAGCTGGAATCTCTTTGCAATCCCATCTCTATAAGATTAAAAAGTTCTCTCAGCGAGTCCC 1150  
 Db 853 CAGCTGGAATCTCTTTGCAATCCCATCTCTATAAGATTAAAAAGTTCTCTCAGCGAGTCCC 912  
 Qy 1151 ATGACCGACAGCCAAAGAAAGCCATCACCAGAAAGATGATATTTGGTATCCGAATGA 1210  
 Db 913 ATGACCGACAGCCAAAGAAAGCCATCACCAGAAAGATGATATTTGGTATCCGAATGA 972  
 Qy 1211 TGCTACTCAGCTCCAGCAAGCAAAAGAGCTGAAGACCGGCGCCGCGGACTTCCAGACG 1270  
 Db 973 TGCTACTCAGCTCCAGCAAGCAAAAGAGCTGAAGACCGGCGCCGCGGACTTCCAGACG 1032  
 Qy 1271 TGATCTCAGGAGCGTATATAATTGAAGTAACTCTGATACCCAGCAGAAAGCTGGTGTC 1330  
 Db 1033 TGATCTCAGGAGCGTATATAATTGAAGTAACTCTGATACCCAGCAGAAAGCTGGTGTC 1092  
 Qy 1331 TCAAGGAAACAGCTCATATAATCAGCATCAATGACAGTCCGTGGTCTCCGCCAATGATG 1390  
 Db 1093 TCAAGGAAACAGCTCATATAATCAGCATCAATGACAGTCCGTGGTCTCCGCCAATGATG 1152  
 Qy 1391 TCAGCGAGCTCATATAAGGGAAGCACCCTGAACATGGTGGTCCGCAAGGCGTAATGAAG 1450  
 Db 1153 TCAGCGAGCTCATATAAGGGAAGCACCCTGAACATGGTGGTCCGCAAGGCGTAATGAAG 1211  
 Qy 1451 ATATCATGATCAGTGATTCGCGAGAAATTCACCCATAGGAGGAGGATGAGCTGGAC 1510  
 Db 1212 ATATCATGATCAGTGATTCGCGAGAAATTCACCCATAGGAGGAGGATGAGCTGGAC 1271  
 Qy 1511 TTCAATGTTCCCTCAAAAGACTCTCCGTTGATGACCGGATGAGGACTCTGGGCTGCGAA 1570  
 Db 1272 TTCAATGTTCCCTCAAAAGACTCTCCGTTGATGACCGGATGAGGACTCTGGGCTGCGAA 1331  
 Qy 1571 TAGGACACTCAAGACTTTTGGTACTGCAATTTTGTTCAGTGAGACTCCCTGGCCAAAC 1630  
 Db 1332 TAGGACACTCAAGACTTTTGGTACTGCAATTTTGTTCAGTGAGACTCCCTGGCCAAAC 1391  
 Qy 1631 AGAATCCTCTTGTATGATTTGAGGCAAAACAAATGTAATGTTGCAGATCCGCGAGGAGA 1690  
 Db 1392 AGAATCCTCTTGTATGATTTGAGGCAAAACAAATGTAATGTTGCAGATCCGCGAGGAGA 1451  
 Qy 1691 AGCTCTGCCCTCTGTATCTATGATGAGTGTGCTTTTTCTTCCAGCTTTGGGCCATT 1750  
 Db 1452 AGCTCTGCCCTCTGTATCTATGATGAGTGTGCTTTTTCTTCCAGCTTTGGGCCATT 1511  
 Qy 1751 CTTGCTTAGACAGTACGATTTGCTCTCTCTTAATGATGATCATCTTAGTCCCACT 1810

Db 1512 CTTGCTTAGACAGTCAGCATTTGTCTCTCTTTTAACCTGAGTCATCATCTTAGTCCAAC 1571  
 Qy 1811 AATGCACTCGATACAAATGCGTAGATAGAAAGCCCAAGGAGCCAGGATGGACTGGT 1870  
 Db 1572 AATGCACTCGATACAAATGCGTAGATAGAAAGCCCAAGGAGCCAGGATGGACTGGT 1631  
 Qy 1871 CGTGTGCTGCTTTCTTCAAGTCAGCACCCAAAGGTCAATGACAGAGACCCCGGTGG 1930  
 Db 1632 CGTGTGCTGCTTTCTTCAAGTCAGCACCCAAAGGTCAATGACAGAGACCCCGGTGG 1691  
 Qy 1931 GTGAGCGCTGGCTTCTCAAAACGCGCAAGTGTGCTCTTTTAGGAATCTCTTTGGAATTGG 1990  
 Db 1692 GTGAGCGCTGGCTTCTCAAAACGCGCAAGTGTGCTCTTTTAGGAATCTCTTTGGAATTGG 1751  
 Qy 1991 GAGCACGATGACTCTGAGTTTGAAGTATTAAGTACTTCTTACATTTGAATAAAAAA 2050  
 Db 1752 GAGCACGATGACTCTGAGTTTGAAGTATTAAGTACTTCTTACATTTGAATAAAAAA 1811  
 Qy 2051 AAAA 2054  
 Db 1812 AAAA 1815

RESULT 7  
 US-10-104-047-795  
 ; Sequence 795, Application US/10104047  
 ; Publication No. US20030236392A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: HELIX RESEARCH INSTITUTE  
 ; TITLE OF INVENTION: No. US20030236392A1e1 full length cDNA  
 ; FILE REFERENCE: H1-A0105  
 ; CURRENT APPLICATION NUMBER: US/10/104,047  
 ; PRIOR FILING DATE: 2002-03-25  
 ; PRIOR APPLICATION NUMBER:  
 ; NUMBER OF SEQ ID NOS: 4096  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 795  
 ; LENGTH: 1894  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-10-104-047-795

Query Match 82.8%; Score 1701.4; DB 15; Length 1894;  
 Best Local Similarity 97.2%; Pred. No. 0;  
 Matches 1761; Conservative 0; Mismatches 1; Indels 49; Gaps 1;

Qy 229 GCGTCCGCTGTCGAGAGTGTGCGGCGCCCGAGGCGCGGCGTGGCGCTTCGAGGAG 288  
 Db 133 GTGTGCGGCTGTCGAGAGTGTGCGGCGCCCGAGGCGCGGCGTGGCGCTTCGAGGAG 192  
 Qy 289 GCGCCGTGCGGCGAGGCGTGCAGTGTGTCGCTTCCGCGGTCGCCAGCCTCGGCCACG 348  
 Db 193 GCGCCGTGCGGCGAGGCGTGCAGTGTGTCGCTTCCGCGGTCGCCAGCCTCGGCCACG 252  
 Qy 349 GTGCGGCGGCGCGAGGCGGCTTCTGTGTGTGCGGCGAGCAGCGCGGTGTGCGCG 408  
 Db 253 GTGCGGCGGCGCGAGGCGGCTTCTGTGTGTGCGGCGGCTTCTGTGTGCGGCGG 293  
 Qy 409 AGCGAGCCCAACACCTACGCCACCTGTGTCAGCTGCGCGCGCCAGCGCGCGCTCCGAG 468  
 Db 294 -----CAGTGTGCGGCGCGCGCGCGCGCGCTCCGAG 323  
 Qy 469 AGCGTGCACCGCGCGCGGTCATCGCTCAGCGCGGAGCTCGCGGCAAGGGCAGGAA 528  
 Db 324 AGCGTGCACCGCGCGCGGTCATCGCTCAGCGCGGAGGCTCGCGGCAAGGGCAGGAA 383  
 Qy 529 GATCCCAACAGTTTGGCCCAATAATACTTTATCGGAGAGCTGGTGGAGAAAGTTCGCC 588  
 Db 384 GATCCCAACAGTTTGGCCCAATAATACTTTATCGCGGAGCTGGTGGAGAAAGTTCGCC 443  
 Qy 589 CTTGCGCTGCTCATATCGAATTTGTTTCGCAAGCTTCCGTTTCTTAAACGAGAGGTGCGG 648

Db 444 CTTGCGGTGGTTTCATATCGAATGTTTTCGCAAGCTTCGGTTTCTAAACGAGAGGTGCGG 503  
Qy 649 GTGGCTAGTGGGTCTGGTTTATTTGTGTGCGAAGATGGAATGATCGTGACAAATGCCAC 708  
Db 504 GTGGCTAGTGGGTCTGGTTTATTTGTGTGCGAAGATGGAATGATCGTGACAAATGCCAC 563  
Qy 709 GTGGTGACCAACAAAGCACCGGGTCAAAAGTTGAGCTGAAGAACGGTGGCCATTTACGAAGCC 768  
Db 564 GTGGTGACCAACAAAGCACCGGGTCAAAAGTTGAGCTGAAGAACGGTGGCCATTTACGAAGCC 623  
Qy 769 AAAATCAAGGATGTGATGAGAAAGCAGACATCGCACTCATCAAAATTTGACCACCAAGGC 828  
Db 624 AAAATCAAGGATGTGATGAGAAAGCAGACATCGCACTCATCAAAATTTGACCACCAAGGC 683  
Qy 829 AAGCTCCCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCT 888  
Db 684 AAGCTCCCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCT 743  
Qy 889 GCATCGGAAGCCCGTTTCCCTTCAAAACACAGTCAACCACCGGATCGTGACCAACC 948  
Db 744 GCATCGGAAGCCCGTTTCCCTTCAAAACACAGTCAACCACCGGATCGTGACCAACC 803  
Qy 949 CAGCAGGGCGCAAGAGCTGGGGCTCCGCAACTCAGACATGGAATCATCCAGACCGAC 1008  
Db 804 CAGCAGGGCGCAAGAGCTGGGGCTCCGCAACTCAGACATGGAATCATCCAGACCGAC 863  
Qy 1009 GCATCATCAACTATGGAATCTGGGAGCCGCTAGTAAACCTGGAGCGTGAAGTATT 1068  
Db 864 GCATCATCAACTATGGAATCTGGGAGCCGCTAGTAAACCTGGAGCGTGAAGTATT 923  
Qy 1069 GGAATTAACACTTTGAAAGTGACAGCTGGAATCTCTTTGCAATCCCATCTGATAAGATT 1128  
Db 924 GGAATTAACACTTTGAAAGTGACAGCTGGAATCTCTTTGCAATCCCATCTGATAAGATT 983  
Qy 1129 AAAAGTCTCTACGAGTCCCATGACGAGCCAGCCAAAGGAAAGCCATCACCAAGAG 1188  
Db 984 AAAAGTCTCTACGAGTCCCATGACGAGCCAGCCAAAGGAAAGCCATCACCAAGAG 1043  
Qy 1189 AAGTATATTGGTATCGAATGATGTCTACTCAGCTCCAGCAAGCCAAAGAGCTGAAGAC 1248  
Db 1044 AAGTATATTGGTATCGAATGATGTCTACTCAGCTCCAGCAAGCCAAAGAGCTGAAGAC 1103  
Qy 1249 CGGCACCGGACTTCCAGAGCTGATCTCAGGAGGTATATAATTGAAGTAACTCTGAT 1308  
Db 1104 CGGCACCGGACTTCCAGAGCTGATCTCAGGAGGTATATAATTGAAGTAACTCTGAT 1163  
Qy 1309 ACCCCAGCAGAGCTGGTGTCTCAAGGAAACGACGTCATAATCAGCATCAATGGACAG 1368  
Db 1164 ACCCCAGCAGAGCTGGTGTCTCAAGGAAACGACGTCATAATCAGCATCAATGGACAG 1223  
Qy 1369 TCGGTGCTCCGCAATGATGTCAGGAGCTCATTAAGGGAAGCACCTGACATG 1428  
Db 1224 TCGGTGCTCCGCAATGATGTCAGGAGCTCATTAAGGGAAGCACCTGACATG 1283  
Qy 1429 GTGGTCCGAGGGGTAAAGAATATCATGATCAGTGATTTCCCGAAGAAATTGACCCA 1488  
Db 1284 GTGGTCCGAGGGGTAAAGAATATCATGATCAGTGATTTCCCGAAGAAATTGACCCA 1343  
Qy 1489 TAGGCAGAGCATGAGTGCATTCATGTTTCCCTCAAGACTCTCCGCTGATACGGA 1548  
Db 1344 TAGGCAGAGCATGAGTGCATTCATGTTTCCCTCAAGACTCTCCGCTGATACGGA 1403  
Qy 1549 TGAGGACTCTGGGTGCTGGAATAGGACACTCAAGACTTTTACTGCCATTTTGTGTT 1608  
Db 1404 TGAGGACTCTGGGTGCTGGAATAGGACACTCAAGACTTTTACTGCCATTTTGTGTT 1463  
Qy 1609 CAGTGAGACTCTCGGCAACAGAACTCTTCTGATAGTTTGGAGGCAAAACAATGTA 1668  
Db 1464 CAGTGAGACTCTCGGCAACAGAACTCTTCTGATAGTTTGGAGGCAAAACAATGTA 1523  
Qy 1669 ATGTTGAGATCCGACGAGAACTCTGCCCTTCTGTATCCTATGATGAGTGTGCTT 1728  
Db 1524 ATGTTGAGATCCGACGAGAACTCTGCCCTTCTGTATCCTATGATGAGTGTGCTT 1583

Qy 1729 TTTCTTCCAGCTTGGGCCATTCTTCTTACACAGTCAGCATTTGTCTCTCTCTTAACT 1788  
Db 1584 TTTCTTCCAGCTTGGGCCATTCTTCTTACACAGTCAGCATTTGTCTCTCTCTTAACT 1643  
Qy 1789 GAGTCATCATCTTTAGTCCAACTAATGCACTGAGTCGATCAATGCGTAGAAGAAGCCCCA 1848  
Db 1644 GAGTCATCATCTTTAGTCCAACTAATGCACTGAGTCGATCAATGCGTAGAAGAAGCCCCA 1703  
Qy 1849 CGGAGCCAGATGGGACTGTGCTGTTTGTGCTTTCTCCAAAGTCAGCACCCAAAGGTC 1908  
Db 1704 CGGAGCCAGATGGGACTGTGCTGTTTGTGCTTTCTCCAAAGTCAGCACCCAAAGGTC 1763  
Qy 1909 AATGCAAGAGACCCCGGTGGGTGAGCGCTGGCTTCTCAAACGGCCGAAGTTGCCTTT 1968  
Db 1764 AATGCAAGAGACCCCGGTGGGTGAGCGCTGGCTTCTCAAACGGCCGAAGTTGCCTTT 1823  
Qy 1969 TTAGGAATCTTTTGAATTTGGGAGCAGATGACTCTGAGTTTGTAGCTATTAAAGTACTT 2028  
Db 1824 TTAGGAATCTTTTGAATTTGGGAGCAGATGACTCTGAGTTTGTAGCTATTAAAGTACTT 1883  
Qy 2029 CTTACACATTG 2039  
Db 1884 CTTACACATTG 1894

## RESULT 8

US-09-765-231A-20  
; Sequence 20, Application US/09765231A  
; Patent No. US20020119452A1  
; GENERAL INFORMATION:  
; APPLICANT: Searle/Monsanto  
; APPLICANT: Phippard, Deborah  
; APPLICANT: Vasanthakumar, Geetha  
; APPLICANT: Dotson, Stanton  
; APPLICANT: Ma, Xiao-Jun  
; TITLE OF INVENTION: Osteoarthritis tissue-derived nucleic acids, polypeptides,  
; TITLE OF INVENTION: vectors, and cells  
; FILE REFERENCE: SO-3221 PR  
; CURRENT APPLICATION NUMBER: US/09/765,231A  
; CURRENT FILING DATE: 2001-01-18  
; NUMBER OF SEQ ID NOS: 82  
; SEQ ID NO 20  
; LENGTH: 1534  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-765-231A-20

Query Match 70.8%; Score 1455.2; DB 9; Length 1534;  
Best Local Similarity 99.3%; Pred. No. 0;  
Matches 1524; Conservative 0; Mismatches 3; Indels 7; Gaps 6;

Qy 525 GGAAGATCCCAACAGTTTGGCCATAAATAAATTAACATTTATCGGAGCGTGTGGAGAAGAT 584  
Db 1 GGAAGATCCCAACAGTTTGGCCATAAATAAATTAACATTTATCGGAGCGTGTGGAGAAGAT 60  
Qy 585 CGCCCTCGCGTGGTTCATATCGAATTTGCGCAAGCTTCCGTTTCTTAAACGAGAGGT 644  
Db 61 CGCCCTCGCGTGGTTCATATCGAATTTGCGCAAGCTTCCGTTTCTTAAACGAGAGGT 120  
Qy 645 GCGGTGGCTAGTGGGTCTGGGTTTATGTCGGAAGATGGAATGATCGTGAACAATGC 704  
Db 121 GCGGTGGCTAGTGGGTCTGGGTTTATGTCGGAAGATGGAATGATCGTGAACAATGC 180  
Qy 705 CCACGTGGTGACCAACAAAGCACCCGGTCAAAGTTGAGCTGAAGAAAGCGTGCCACTTACGA 764  
Db 181 CCACGTGGTGACCAACAAAGCACCCGGTCAAAGTTGAGCTGAAGAAAGCGTGCCACTTACGA 240  
Qy 765 AGCCAAATCAAGATGTGATGAAAGCAGACATCGCACTCATCAAAATGACACCA 824  
Db 241 AGCCAAATCAAGATGTGATGAAAGCAGACATCGCACTCATCAAAATGACACCA 300  
Qy 825 GGGCAAGCTGCCTGCTCTGCTGCTTGGCGGCTCCTCAGAGCTGCGGCCGGAGAGTTCGT 884



Db 61 TCGGCGCAGCTGTCCGGGCGCGCGCTTGGCGCCTTGGCGCGCGGTGCCAGACGC 120  
 Qy 169 TCGAGCGCGCGCGCTGCGCGCGCAGCGAGCACTGCGAGGGCGCGCGCGCGCGGAC 228  
 Db 121 TCGAGCGCGCGCGCTGCGCGCGCAGCGAGCACTGCGAGGGCGCGCGCGCGGAC 180  
 Qy 229 GCGTGGCGTGTCTGCGAGGTGTGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 288  
 Db 181 GCGTGGCGTGTCTGCGAGGTGTGCGGCGCGCGCGCGCGCGCGCGCGCGCGCG 240  
 Qy 289 GCGCGTGGCGCGAGGCGTGTGCGAGGTGTGCGGCGCGCGCGCGCGCGCGCGCGCG 348  
 Db 241 GCGCGTGGCGCGAGGCGTGTGCGAGGTGTGCGGCGCGCGCGCGCGCGCGCGCGCG 300  
 Qy 349 GTGCG 408  
 Db 301 GTGCG 360  
 Qy 409 AGCGAGCGCAACCTTACGCGCACTGTGCGAGCTGTGCGCGCGCGCGCGCGCGCGCG 468  
 Db 361 AGCGAGCGCAACCTTACGCGCACTGTGCGAGCTGTGCGCGCGCGCGCGCGCGCGCG 420  
 Qy 469 AGGCTCACCG 528  
 Db 421 AGGCTCACCG 480  
 Qy 529 GATCCCAACAGTTTGGCGCAATAATACTTATCGGAGCGTGTGCGAGAGTATCGCC 588  
 Db 481 GATCCCAACAGTTTGGCGCAATAATACTTATCGGAGCGTGTGCGAGAGTATCGCC 540  
 Qy 589 CTGCGCGTGTCTCATATCGAATTTGTTGCGAAGCTTCTGCGAGCGTGTGCGAGAGTGTGCG 648  
 Db 541 CTGCGCGTGTCTCATATCGAATTTGTTGCGAAGCTTCTGCGAGCGTGTGCGAGAGTGTGCG 600  
 Qy 649 GTGCGTAGTGGTGTGCGGTTTATGTCGGAAGTGTGCGAGAGTGTGCGAGAGTGTGCG 708  
 Db 601 GTGCGTAGTGGTGTGCGGTTTATGTCGGAAGTGTGCGAGAGTGTGCGAGAGTGTGCG 660  
 Qy 709 GTGCGTAGTGGTGTGCGGTTTATGTCGGAAGTGTGCGAGAGTGTGCGAGAGTGTGCG 768  
 Db 661 GTGCGTAGTGGTGTGCGGTTTATGTCGGAAGTGTGCGAGAGTGTGCGAGAGTGTGCG 720  
 Qy 769 AAAATCAAGATGTGATGAGAGAGAGATGCGACTCATCAAAATGACCAAGCGG 828  
 Db 721 AAAATCAAGATGTGATGAGAGAGAGATGCGACTCATCAAAATGACCAAGCGG 780  
 Qy 829 AAGCTCCCTGTCTGTCTGTGCGCGCTTCTCAGAGCTGCGCGCGGAGAGTGTGCGTGC 888  
 Db 781 AAGCTCCCTGTCTGTCTGTGCGCGCTTCTCAGAGCTGCGCGCGGAGAGTGTGCGTGC 840  
 Qy 889 GGCATCGAAGCGCGTTTCCCTTCAAAACACAGTCAACCGGATCGTGAGCAACACC 948  
 Db 841 GGCATCGAAGCGCGTTTCCCTTCAAAACACAGTCAACCGGATCGTGAGCAACACC 900  
 Qy 949 CAGCAGCGCGGCAAGAGCTGGGCTCGCAACTCAGACATGGAATCATCCAGACCGAC 1008  
 Db 901 CAGCAGCGCGGCAAGAGCTGGGCTCGCAACTCAGACATGGAATCATCCAGACCGAC 960  
 Qy 1009 GGCATCATCAACTATGGAATCTGGGAGCGCGTTAGTAAACCTGAGCGGTGAAGTAT 1068  
 Db 961 GGCATCATCAACTATGGAATCTGGGAGCGCGTTAGTAAACCTGAGCGGTGAAGTAT 1020  
 Qy 1069 GGAATTAACACTTTGAAAGTGACAGCTGGAATCTCTTTGCAATCCCATCTGATAGATT 1128  
 Db 1021 GGAATTAACACTTTGAAAGTGACAGCTGGAATCTCTTTGCAATCCCATCTGATAGATT 1080  
 Qy 1129 AAAAGTTCTCAGGAGTCCCATGACGACAGCGCAAGAGAGAGAGAGAGAGAGAGAG 1188  
 Db 1081 AAAAGTTCTCAGGAGTCCCATGACGACAGCGCAAGAGAGAGAGAGAGAGAGAGAG 1140  
 Qy 1189 AAGTATATGCTATCCGAATGATGCTACCTACGCTCCAGCAAGAGAGAGAGAGAGAG 1248  
 Db 1141 AAGTATATGCTATCCGAATGATGCTACCTACGCTCCAGCAAGAGAGAGAGAGAGAG 1200

RESULT 10

US-10-198-846-14051  
 ; Sequence 14051, Application US/10198846  
 ; Publication No. US2003009974A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Lillie, James  
 ; APPLICANT: Xu, Yongyao  
 ; APPLICANT: Wang, Youzhen  
 ; APPLICANT: Steinmann, Kathleen  
 ; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS  
 ; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND  
 ; TITLE OF INVENTION: THERAPY OF BREAST CANCER  
 ; FILE REFERENCE: MRI-049  
 ; CURRENT APPLICATION NUMBER: US/10/198,846  
 ; CURRENT FILING DATE: 2002-07-18  
 ; PRIOR APPLICATION NUMBER: 60/306,220  
 ; PRIOR FILING DATE: 2001-07-18  
 ; NUMBER OF SEQ ID NOS: 14084  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 14051  
 ; LENGTH: 1498  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-10-198-846-14051

Query Match 62.5%; Score 1283.4; DB 14; Length 1498;  
 Best Local Similarity 91.4%; Pred. No. 0;  
 Matches 1427; Conservative 0; Mismatches 1; Indels 133; Gaps 1;

Qy 479 GGCGCGCGGTCTCTGTCAGCGCGAGCTGCGGCAAGCGGAGAGATCCCAACA 538  
 Db 71 GGCGCGCGGTCTCTGTCAGCGCGAGCTGCGGCAAGCGGAGAGATCCCAACA 130  
 Qy 539 GTTTGGCGCAATAATAAATCTTTATCGCGGAGCTGTGAGAGAGATCGCCCTGCGGTGG 598  
 Db 131 GTTTGGCGCAATAATAAATCTTTATCGCGGAGCTGTGAGAGAGATCGCCCTGCGGTGG 190  
 Qy 599 TCCATATCGAATTTGTTTCGCAAGCTTCTTAAACGAGAGGTGCGGTGGCTAGTG 658  
 Db 191 TCCATATCGAATTTGTTTCGCAAGCTTCTTAAACGAGAGGTGCGGTGGCTAGTG 250  
 Qy 659 GGTCTGGTTTATGTCGGAAGATGACATGATCGTGAATAATGCCACGTGGTGAACA 718  
 Db 251 GGTCTGGTTTATGTCGGAAGATGACATGATCGTGAATAATGCCACGTGGTGAACA 310  
 Qy 719 ACAAGCACCGGTCAGAGTTGAGCTGAAGAGCGGTGCCACTTACGAAGCCAAATCAAGG 778  
 Db 311 ACAAGCACCGGTCAGAGTTGAGCTGAAGAGCGGTGCCACTTAC----- 354  
 Qy 779 ATGTGGATGAGAAAGCAGACATCGCACTCATCAAAATTTGACCACCGAGGCAAGCTGCTG 838  
 Db 355 ----- 354  
 Qy 839 TCCTGTGCTTGGCGCTCCTCCTCAGAGCTGCGCCGCGAGAGTTCTGTGTCGCCATCGAA 898



355	----	GA 357
Db		
Qy	999	GCCTGTTTCCCTTCAAAACACAGTCAACCACCGGATCGTGAGCACCAACCGAGCGGCG 958
Db	358	GCCTGTTTCCCTTCAAAACACAGTCAACCACCGGATCGTGAGCACCAACCGAGCGGCG 417
Qy	959	GCAAGAGCTGGGCTCCGCAACTCAGACATGGACTACATCAGACACCGACGCATCATCA 1018
Db	418	GCAAGAGCTGGGCTCCGCAACTCAGACATGGACTACATCAGACACCGACGCATCATCA 477
Qy	1019	ACTATGGAACCTCGGAGGCCGTTAGTAAACCTGGACGGTGAAGTGAATTGGAATTAA 1078
Db	478	ACTATGGAACCTCGGAGGCCGTTAGTAAACCTGGACGGTGAAGTGAATTGGAATTAA 537
Qy	1079	CTTTGAAAGTGACAGCTGGAATCTCTTTGCAATCCCAATCTGATAAGATTTAAAAGTTCC 1138
Db	538	CTTTGAAAGTGACAGCTGGAATCTCTTTGCAATCCCAATCTGATAAGATTTAAAAGTTCC 597
Qy	1139	TCACGGAGTCCCATGACCGACAGGCCAAAGAAAGCCATCACCAAGAAAGATGATTTG 1198
Db	598	TCACGGAGTCCCATGACCGACAGGCCAAAGAAAGCCATCACCAAGAAAGATGATTTG 657
Qy	1199	GTATCCGAATGATGTCACTCAGCTCCAGCAAAAGCCAAAGAGCTGAAGACCCGACCGG 1258
Db	658	GTATCCGAATGATGTCACTCAGCTCCAGCAAAAGCCAAAGAGCTGAAGACCCGACCGG 717
Qy	1259	ACTTCCAGACGTGATCTCAGGAGCGTATATAATTGAAGTAATTCCTGATACCCAGCAG 1318
Db	718	ACTTCCAGACGTGATCTCAGGAGCGTATATAATTGAAGTAATTCCTGATACCCAGCAG 777
Qy	1319	AAGCTGGTGGTCTCAAGGAAAAAGACGCTCAATAACAGCATCAATGGACAGTCCGTGGTCT 1378
Db	778	AAGCTGGTGGTCTCAAGGAAAAAGACGCTCAATAACAGCATCAATGGACAGTCCGTGGTCT 837
Qy	1379	CGCCAAATGATGTACAGGAGCTCATTAAGAGGAAAGACACCTGAACATGGTGGTCCGCA 1438
Db	838	CGCCAAATGATGTACAGGAGCTCATTAAGAGGAAAGACACCTGAACATGGTGGTCCGCG 897
Qy	1439	GGGGTAATGAAGATATCATGATCACAGTGAATCCCGAAGAAATTCACCCATAGGCAGAGG 1498
Db	898	GGGGTAATGAAGATATCATGATCACAGTGAATCCCGAAGAAATTCACCCATAGGCAGAGG 957
Qy	1499	CATGAGCTGGACTTCATGTTTCCTCAAAGACTCTCCCGTGGATGACGGATGAGGACTCT 1558
Db	958	CATGAGCTGGACTTCATGTTTCCTCAAAGACTCTCCCGTGGATGACGGATGAGGACTCT 1017
Qy	1559	GGGCTGCTGGAAATAGACACTCAAGACTTTTGAAGTGCATGCCATTTGTTTGTTCAGTGAGAC 1618
Db	1018	GGGCTGCTGGAAATAGACACTCAAGACTTTTGAAGTGCATGCCATTTGTTTGTTCAGTGAGAC 1077
Qy	1619	TCCCTGGCCAAAGAAATCCTTCTTGATAGTATTGACGGCAAAACAAATGTAATGTTCCAGA 1678
Db	1078	TCCCTGGCCAAAGAAATCCTTCTTGATAGTATTGACGGCAAAACAAATGTAATGTTCCAGA 1137
Qy	1679	TCCGAGGCAGAGACTCTGCCTTCTGTATCCTATGATGCAAGTGTGCTTTTTCTTTGCCA 1738
Db	1138	TCCGAGGCAGAGACTCTGCCTTCTGTATCCTATGATGCAAGTGTGCTTTTTCTTTGCCA 1197
Qy	1739	GCTTGGGCCATTCTTGCTTAGACAGTCAGCAATTTGTCTCCTCTTTAACTGAGTCATCAT 1798
Db	1198	GCTTGGGCCATTCTTGCTTAGACAGTCAGCAATTTGTCTCCTCTTTAACTGAGTCATCAT 1257
Qy	1799	CTTAGTCCAACTAATGCAGTCGATACAAATGGCTAGATAGAAGAGCCCAACCGGAGCCAG 1858
Db	1258	CTTAGTCCAACTAATGCAGTCGATACAAATGGCTAGATAGAAGAGCCCAACCGGAGCCAG 1317
Qy	1859	GATGGGACTGGTCTGCTTTGCTTTCTCCAAAGTCAGACACCCAAAGGTCATATGCACAGA 1918
Db	1318	GATGGGACTGGTCTGCTTTGCTTTCTCCAAAGTCAGACACCCAAAGGTCATATGCACAGA 1377
Qy	1919	GACCCGGGTGGTGAGCGTGGCTTCTCAAACCGCCGAGTTGCTCTTTTAGGAATCT 1978
Db	1378	GACCCGGGTGGTGAGCGTGGCTTCTCAAACCGCCGAGTTGCTCTTTTAGGAATCT 1437

Qy	1979	CTTTGGAAATTGGGAGCAGCATGACTCTGAGTTTGAGCTATTAAAGTACTTCTTACACATT	2038
Db	1438	CTTTGGAAATTGGGAGCAGCATGACTCTGAGTTTGAGCTATTAAAGTACTTCTTACACATT	1497
Qy	2039	G 2039	
Db	1498	G 1498	
RESULT 11			
US-09-974-298-67			
; Sequence 67, Application US/09974298			
; Patent No. US20020156263A1			
; GENERAL INFORMATION:			
; APPLICANT: Chen, Hwei-Mei			
; TITLE OF INVENTION: GENES EXPRESSED IN BREAST CANCER			
; FILE REFERENCE: PA-0037 P			
; CURRENT APPLICATION NUMBER: US/09/974,298			
; CURRENT FILING DATE: 2001-10-04			
; PRIOR APPLICATION NUMBER: 60/238,331			
; PRIOR FILING DATE: 2000-05-10			
; NUMBER OF SEQ ID NOS: 194			
; SOFTWARE: PERL Program			
; SEQ ID NO 67			
; LENGTH: 1389			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
; FEATURE:			
; NAME/KEY: misc_feature			
; OTHER INFORMATION: Incyte ID No. US20020156263A1 1100429.10			
; NAME/KEY: unsure			
; LOCATION: 1372			
; OTHER INFORMATION: a, t, c, g, or other			
US-09-974-298-67			
Query Match 37.0%; Score 759.8; DB 9; Length 1389;			
Best Local Similarity 99.5%; Pred. No. 1e-204;			
Matches 772; Conservative 0; Mismatches 3; Indels 1; Gaps 1;			
Qy	1219	ACGTCACGAAGCCAAAGAGCTGAAGGACGGGACCGGGACTTCCCAGACGTGATCTCA	1278
Db	615	ACATCCAGCAAGCCAAAGAGCTGAAGGACGGGACCGGGACTTCCCAGACGTGATCTCA	674
Qy	1279	GGAGCGGTATATAATTGAAGTAAATTCCTGTATACCCACAGAAAGCTGGTGGTCTCAAGGAA	1338
Db	675	GGAGCGGTATATAATTGAAGTAAATTCCTGTATACCCACAGAAAGCTGGTGGTCTCAAGGAA	734
Qy	1339	AACGAGGTCATATACGATCAATGNCAGTCCGTGGTCTCGGCCAAATGATGTACGGAC	1398
Db	735	AACGAGGTCATATACGATCAATGNCAGTCCGTGGTCTCGGCCAAATGATGTACGGAC	794
Qy	1399	GTCAATTAAGGGAAAGACCCCTGAACATGGTGGTCCGACGGGGTAAATGAAGATATCATG	1458
Db	795	GTCAATTAAGGGAAAGACCCCTGAACATGGTGGTCCGACGGGGTAAATGAAGATATCATG	854
Qy	1459	ATCAGAGTATTTCCCGAAGAAATTGACCATAGGACGGGATGAGCTGGACTTTCATGTT	1518
Db	855	ATCAGAGTATTTCCCGAAGAAATTGACCATAGGACGGGATGAGCTGGACTTTCATGTT	914
Qy	1519	TCCCTCAAGACTCTCCCGTGGATAGCGGATAGGACTCTGGGCTGTGGATAGGACAC	1578
Db	915	TCTCTCAAGACTCTCCCGTGGATAGCGGATAGGACTCTGGGCTGTGGATAGGACAC	974
Qy	1579	TCAAGACTTTTGACTGCCATTTTGTGTTTGTTCAGTGGAGACTCCCTGGGCCAACAGAAATCCT	1638
Db	975	TCAAGACTTTTGACTGCCATTTTGTGTTTGTTCAGTGGAGACTCCCTGGGCCAACAGAAATCCT	1034
Qy	1639	TCTTGATGTTGACGGCAAAACAAATGTAATGTTGCAGATCCGACGAGAGCTCTGC	1698
Db	1035	TCTTGATAG-TTGACGGCAAAACAAATGTAATGTTGCAGATCCGACGAGAGCTCTGC	1093
Qy	1699	CCTTCTGTATCTTATGTATGCAGTGTGCTTTTCTTTCGACGCTTGGGCCATCTTCTGCTTA	1758

Query Match	37.0%;	Score 759.8;	DB 9;	Length 1389;
Best Local Similarity	99.5%;	Pred. No. 1e-204;		
Matches 772;	Conservative 0;	Mismatches 3;	Indels 1;	Gaps 1;
QY	1219	ACGTCAGCAAAAGCCAAAGAGCGCTGAAGGACCGGCACCGGGACTTCCCGAGACGTGATCTCA	1278	
DB	615	ACATCCAGCAAAAGCCAAAGAGCGCTGAAGGACCGGCACCGGGACTTCCCGAGACGTGATCTCA	674	
QY	1279	GGAGCGGTATATAATTGAAGTAATTCCTGATACCCGACGAGAAGCTGTGGTCTCAAGGA	1338	
DB	675	GGAGCGGTATATAATTGAAGTAATTCCTGATACCCGACGAGAAGCTGTGGTCTCAAGGA	734	
QY	1339	AACGACGTCATAATCAGCATCAATGGACAGCTCCGTGGTCTCGGCCAATGATGTCACGCAC	1398	
DB	735	AACGACGTCATAATCAGCATCAATGGACAGCTCCGTGGTCTCGGCCAATGATGTCACGCAC	794	
QY	1399	GTCATTAAGGGAAGACACCCCTGAACATGGTGTCTCGCAGGGTAAATGAAGATATCATG	1458	
DB	795	GTCATTAAGGGAAGACACCCCTGAACATGGTGTCTCGCAGGGTAAATGAAGATATCATG	854	
QY	1459	ATCACAGTGATTCCCGAAGAAATTGACCCATAGGCAGAGGATGAGCTGGACTTCATGTT	1518	
DB	855	ATCACAGTGATTCCCGAAGAAATTGACCCATAGGCAGAGGATGAGCTGGACTTCATGTT	914	
QY	1519	TCCCTCAAAGACTCTCCGTGGATGACGGATGAGGACTCTGGGCTGCTGGATAGACAC	1578	
DB	915	TCTCTCAAAGACTCTCCGTGGATGACGGATGAGGACTCTGGGCTGCTGGATAGACAC	974	
QY	1579	TCAAGACTTTTGACTGCCATTTTGTGTTGTTTCAGTGGAGACTCCCTGGCCAAACAGAAATCCT	1638	
DB	975	TCAAGACTTTTGACTGCCATTTTGTGTTGTTTCAGTGGAGACTCCCTGGCCAAACAGAAATCCT	1034	
QY	1639	TCTTGATAGTTTCAGGCAAAAACAAATGTAATGTTGCAGATCCGACGGCAGAGCTCTGC	1698	
DB	1035	TCTTGATAG-TTGCAGGCAAAAACAAATGTAATGTTGCAGATCCGACGGCAGAGCTCTGC	1093	
QY	1699	CCCTTGATCTCATGTATGCAGGTGTGCTTTTCTTGCCAGCTTGGGCCATTTCTGCTTA	1758	

Db 1094 CTTCTGTAATCCATGATGAGTGTCTTTCTTTCAGAGTTGGGCAATCTTCTGTTA 1153  
 Qy 1759 GACAGTCAGCATTTGTCCTCTCTTAACTCAGTCATCATCTTGTCCAACTAAATGCAAT 1818  
 Db 1154 GACAGTCAGCATTTGTCCTCTCTTAACTCAGTCATCATCTTGTCCAACTAAATGCAAT 1213  
 Qy 1819 CGATACAAATCGGTAGATAGAGAGAGCCCGGAGAGCCAGGATGGAGTGGTCTGTTG 1878  
 Db 1214 CGATACAAATCGGTAGATAGAGAGAGCCCGGAGAGCCAGGATGGAGTGGTCTGTTG 1273  
 Qy 1879 TGCTTTTCTCAAGTCAGCAGCCAAAGAGTCAATGCAGAGACCCCGGTGGGTGAGCGC 1938  
 Db 1274 TGCTTTTCTCAAGTCAGCAGCCAAAGAGTCAATGCAGAGACCCCGGTGGGTGAGCGC 1333  
 Qy 1939 TGCTTTTCTCAAGTCAGCAGCCCGGAGTGGTCTTCTTGAATTTGGGAGC 1994  
 Db 1334 TGCTTTTCTCAAGTCAGCAGCCCGGAGTGGTCTTCTTGAATTTGGGAGC 1389

RESULT 12  
 US-10-198-846-9582  
 ; Sequence 9582, Application US/10198846  
 ; Publication No. US20030099974A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Lillie, James  
 ; APPLICANT: Xu, Yongyao  
 ; APPLICANT: Wang, Youzhen  
 ; APPLICANT: Steinmann, Kathleen  
 ; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS  
 ; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND  
 ; FILE REFERENCE: MEI-049  
 ; CURRENT APPLICATION NUMBER: US/10/198,846  
 ; CURRENT FILING DATE: 2002-07-18  
 ; PRIOR APPLICATION NUMBER: 60/306,220  
 ; PRIOR FILING DATE: 2001-07-18  
 ; NUMBER OF SEQ ID NOS: 14084  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 9582  
 ; LENGTH: 797  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: misc\_feature  
 ; LOCATION: 2, 3, 6, 9, 17, 630, 633, 706, 728, 784, 789  
 ; OTHER INFORMATION: n = A,T,C or G  
 US-10-198-846-9582

Query Match 30.9%; Score 634.4; DB 14; Length 797;  
 Best Local Similarity 96.3%; Pred. No. 2.9e-169;  
 Matches 680; Conservative 0; Mismatches 20; Indels 6; Gaps 3;  
 Qy 1179 CACCAAGAAGATATATTGGTATCCGAATGATGTCACTCAGTCAGCAAGCCAAAGA 1238  
 Db 30 CACCAAGAAGATATATTGGTATCCGAATGATGTCACTCAGTCAGCAAGCCAAAGA 89  
 Qy 1239 GCTGAAGGACCGGACCGGACCTCCAGACGTGATCTCAGGAGGTATATATTTGAAT 1298  
 Db 90 GCTGAAGGACCGGACCGGACCTCCAGACGTGATCTCAGGAGGTATATATTTGAAT 149  
 Qy 1299 AATTCTCTGATACCCAGCAGAGTGGTGTCTCAAGGAAAACGAGTCATATACAGCAT 1358  
 Db 150 AATTCTCTGATACCCAGCAGAGTGGTGTCTCAAGGAAAACGAGTCATATACAGCAT 209  
 Qy 1359 CAATGGACAGTCCGTGGTCTCCGCCAATGATGTACGCGACGTCAATAAAGGAAAGCAC 1418  
 Db 210 CAATGGACAGTCCGTGGTCTCCGCCAATGATGTACGCGACGTCAATAAAGGAAAGCAC 269  
 Qy 1419 CCTGAACATGGTGGTCCGAGGGGTAAAGATATCATGATCAGATATCATGATCAGATGATCCCGAAGA 1478  
 Db 270 CCTGAACATGGTGGTCCGAGGGGTAAAGATATCATGATCAGATGATCCCGAAGA 329

Qy 1479 AATTGACCCATAGCAGAGGCATGAGCTGGACTTTCATGTTTCCCTCAAGACTCTCCCGT 1538  
 Db 330 AATTGACCCATAGCAGAGGCATGAGCTGGACTTTCATGTTTCCCTCAAGACTCTCCCGT 389  
 Qy 1539 GGATGACGGATAGGACTCTGGGCTGTGGAAATAGGACACTCAAGACTTTTGTGATGCAAT 1598  
 Db 390 GGATGACGGATAGGACTCTGGGCTGTGGAAATAGGACACTCAAGACTTTTGTGATGCAAT 449  
 Qy 1599 TTTGTTTGTTCAGTGGAGACTCCCTGGCCACAGAAATCCTTCTTGATAGTTTTCAGGCAA 1658  
 Db 450 TTTGTTTGTTCAGTGGAGACTCCCTGGCCACAGAAATCCTTCTTGATAGTTTTCAGGCAA 509  
 Qy 1659 AACAAATGTATGTGTGAGATCCCGCAGCAGAGCTCTGCCCCTTCTGTATCTTGTATGATG 1718  
 Db 510 AACAAATGTATGTGTGAGATCCCGCAGCAGAGCTCTGCCCCTTCTGTATCTTGTATGATG 569  
 Qy 1719 C-AGTGTGCTTTTCTTGGCCAGC-ITGGGCCAATCTTGTGTTAGACAGTCTGAGCAATTTCTCT 1776  
 Db 570 CAAGTGTGCTTTTCTTGGCCAGCTTTGGGCCAATCTTGTGTTAGACAGTCTGAGCAATTTCTCT 629  
 Qy 1777 CCTCCTTTAACTGAGTCATCATCTTGTAGTCCAACTAATGAGTCGATACATAATGCGTAGAT- 1835  
 Db 630 NCTNCTTTAACTGAGTCATCATCTTGTAGTCCAACTAATGAGTCGATACATAATGCGTAGATA 689  
 Qy 1836 ---AGAGAGAGCCCGGAGCCAGGATGGAGTGGTCTGTTG 1878  
 Db 690 GGAAGAGAGCCCGGAGCCAGGATGGAGTGGTCTGTTG 735

RESULT 13  
 US-10-198-846-9579/c  
 ; Sequence 9579, Application US/10198846  
 ; Publication No. US20030099974A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Lillie, James  
 ; APPLICANT: Xu, Yongyao  
 ; APPLICANT: Wang, Youzhen  
 ; APPLICANT: Steinmann, Kathleen  
 ; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS  
 ; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND  
 ; FILE REFERENCE: MEI-049  
 ; CURRENT APPLICATION NUMBER: US/10/198,846  
 ; CURRENT FILING DATE: 2002-07-18  
 ; PRIOR APPLICATION NUMBER: 60/306,220  
 ; PRIOR FILING DATE: 2001-07-18  
 ; NUMBER OF SEQ ID NOS: 14084  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 9579  
 ; LENGTH: 672  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: misc\_feature  
 ; LOCATION: 14, 21, 23, 27, 29, 30  
 ; OTHER INFORMATION: n = A,T,C or G  
 US-10-198-846-9579

Query Match 30.0%; Score 616.8; DB 14; Length 672;  
 Best Local Similarity 99.7%; Pred. No. 2.7e-164;  
 Matches 618; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 Qy 1337 AAAACGAGCTCATATATCAGCATCAATGGACAGTCCGTGTCTCCGCCAATGATGTACGCG 1396  
 Db 672 AAAACGAGCTCATATATCAGCATCAATGGACAGTCCGTGTCTCCGCCAATGATGTACGCG 613  
 Qy 1397 ACGTCAATAAAAGGAAAGCCCTTGAACATGGTGTCTCCGAGGGGTAAAGATATCA 1456  
 Db 612 ACGTCAATAAAAGGAAAGCCCTTGAACATGGTGTCTCCGAGGGGTAAAGATATCA 553  
 Qy 1457 TGATCAGATGATTTCCCGAAGAAATTGACCCATAGCAGAGGCATGAGCTGGACTTCATG 1516  
 Db 552 TGATCAGATGATTTCCCGAAGAAATTGACCCATAGCAGAGGCATGAGCTGGACTTCATG 493

QY	1517	TTTCCCTCAAAGACTCTCCCGTGGATGACGGATGAGGACTCTGGGTGCTGGAATAGGAC	1576
Db	492	TTTCCCTCAAAGACTCTCCCGTGGATGACGGATGAGGACTCTGGGTGCTGGAATAGGAC	433
QY	1577	ACTCAAGACTTTTGACTGCCATTTCTTTGTTCAGTGGAGACTCCCTGGCCCAACGAATC	1636
Db	432	ACTCAAGACTTTTGACTGCCATTTCTTTGTTCAGTGGAGACTCCCTGGCCCAACGAATC	373
QY	1637	CTTCTTGATAGTTTGAGGGCAAAACAAATGTAAATGTTGCAGATCCGACGACGAAGCTCT	1696
Db	372	CTTCTTGATAGTTTGAGGGCAAAACAAATGTAAATGTTGCAGATCCGACGACGAAGCTCT	313
QY	1697	GCCCTTCGTATCCTATGATGCAGTGTGCTTTTCTTGCCAGCTTGGGCCCATCTCTTGCT	1756
Db	312	GCCCTTCGTATCCTATGATGCAGTGTGCTTTTCTTGCCAGCTTGGGCCCATCTCTTGCT	253
QY	1757	TAGACAGTCAGCAATTTGCTCCTCCTTAACTCAGTCATCATCTTAGTCCCACTAATGCA	1816
Db	252	TAGACAGTCAGCAATTTGCTCCTCCTTAACTCAGTCATCATCTTAGTCCCACTAATGCA	193
QY	1817	GTCGATACAATGCGTAGATAGAAAGACCCACCGGAGCCAGGATGGGACTGGTCGTGTT	1876
Db	192	GTCGATACAATGCGTAGATAGAAAGACCCACCGGAGCCAGGATGGGACTGGTCGTGTT	133
QY	1877	TGTGCTTTTCTCCAGTCAGCACCCAAAGTCAATGCACAGACACCCCGGTGGGTGAGC	1936
Db	132	TGTGCTTTTCTCCAGTCAGCACCCAAAGTCAATGCACAGACACCCCGGTGGGTGAGC	73
QY	1937	GCTGGCTTCTCAAACGGCCG	1956
Db	72	GCTGGCTTATCAAACGGCCG	53

```

RESULT 14
US-10-198-846-13646
; Sequence 13646, Application US/10198846
; Publication No. US20030099974A1
; GENERAL INFORMATION:
; APPLICANT: Lilie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steimmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF BREAST CANCER
; FILE REFERENCE: WRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13646
; LENGTH: 722
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-198-846-13646

```

131	Db	AGCATCAATTGGACAGTCCGTGTCTCCGCCAATGATGTGAGCGACGTCAATTAAGGGAA	190
1414	Qy	AGCACCTGAAACATGGTGTCCGACGGGTAAATGAAGATATCATGATCAAGTGAATCCC	1473
191	Db	AGCACCTGAAACATGGTGTCCGACGGGTAAATGAAGATATCATGATCAAGTGAATCCC	250
1474	Qy	GAAGAAATTGACCCATAGGCACAGGACATGAGCTGGACTTCATGTTTCCCTCAAAGACTCT	1533
251	Db	GAAGAAATTGACCCATAGGCACAGGACATGAGCTGGACTTCATGTTTCCCTCAAAGACTCT	310
1534	Qy	CCGTTGGATGACGGATGAGGACTCTGGGCTGCTGGAAATAGGACACCTCAAGACTTTTGACT	1593
311	Db	CCGTTGGATGACGGATGAGGACTCTGGGCTGCTGGAAATAGGACACCTCAAGACTTTTGACT	370
1594	Qy	GCCATTTGTTTGTTCAGTGGAGACTCCCTGGCCAAACAGAAATCCTCTTGATAGTTTGCA	1653
371	Db	GCCATTTGTTTGTTCAGTGGAGACTCCCTGGCCAAACAGAAATCCTCTTGATAGTTTGCA	430
1654	Qy	GGCAAAACAAATGTAATGTTGAGATCCGACGACAGAAAGCTCTGCGCCTTCTGTATCCTAT	1713
431	Db	GGCAAAACAAATGTAATGTTGAGATCCGCA- GCAGAAAGTTCTGCGCCTTCTGTATCCTAT	489
1714	Qy	GTATGAGTGTGCTTTTCTTGCCAGCTTGGGCCATCTTGCTAGACAGTCAGCATTTG	1773
490	Db	GTATGAGTGTGCTTTTCTTGCCAGCTTGGGCCATCTTGCTTAGAC-----	537
1774	Qy	TCTCCTCCTTAACTGAGTCATCATCTTAGTCCAACTAATGCAAGTCGATACAATGCGTAG	1833
538	Db	-----	537
1834	Qy	ATAGAAGAAGCCCCACGGGAGCCAGGATGGAGTGGTGTGTTTGTGCTTTTCTCCAAAGT	1893
538	Db	-----AGCCAGATGGGACTGGTGGTGTGTTTGTGCTTTTCTCCAAAGT	578
1894	Qy	CAGC-ACCCAAAGGTCATGCAACAGACCCCCGGGTGGGTGAGCGCTGGCTTCTCAAAAG	1952
579	Db	CAGCGACCCAAAGGTCATGCAACAGACCCCCGGGTGGGTGAGCGCTGGCTTCTCAAAAG	638
1953	Qy	GCCGAAGTTGCCTCTTTTAGGAAATCTCTTTGGAAATTTGGGACGACGATGACTCTGAGTTTG	2012
639	Db	GCCGAAGTTGCCTCTTTTAGGAAATCTCTTTGGAAATTTGGGACGACGATGACTCTGAGTTTG	698
2013	Qy	AGCTATTAAGTACTTCTTACACA	2036
699	Db	AGCTATTAAGTACTTCTTACACA	722

```

RESULT 15
US-10-085-783A-14793
; Sequence 14793, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liew, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2002
; CURRENT APPLICATION NUMBER: US/10/085.783A
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 14793
; LENGTH: 575
; TYPE: DNA
; ORGANISM: Human
; US-10-085-783A-14793

```

Query Match 27.2%; Score 559; DB 12; Length 575;

	Best Local Similarity	99.1%;	Pred. No. 6.2e-148;	Matches 562;	Conservative 0;	Mismatches 5;	Indels 0;	Gaps 0.
QY	1488	ATAGGCAGAGGATGAGCTGGACTTCATGTTTCCTCAAAGACTCTCCGCTGGATGACGG	1547					
Db	1	ATAGGCAGAGGATGAGCTGGACTTCATGTTTCCTCAAAGACTCTCCGCTGGATGACGG	60					
QY	1548	ATGAGGACTCTGGGCTGCTGGAAATAGGACACTCAAGACTTTTGACTGCCATTTGTTTGT	1607					
Db	61	ATGAGGACTCTGGGCTGCTGGAAATAGGACACTCAAGACTTTTGACTGCCATTTGTTTGT	120					
QY	1608	TCAGTGGAGACTCCCTGGCCAAACAGAAATCCCTCTTGATAGTTTGCAGGCAAAAACAAATGT	1667					
Db	121	TCAGTGGAGACTCCCTGGCCAAACAGAAATCCCTCTTGATAGTTTGCAGGCAAAAACAAATGT	180					
QY	1668	AATGTTGCAGATCCGGCAGGACAGAGCTCTGCCCTTCTGTATCCTATGTATGCAGTGTGCT	1727					
Db	181	AATGTTGCAGATCCGGCAGGACAGAGCTCTGCCCTTCTGTATCCTATGTATGCAGTGTGCT	240					
QY	1728	TTTTCTTGGCAGCTTGGGCCATCTTGCTTAGACAGTCAGCAATTTGTCTCCTCCTTTAAAC	1787					
Db	241	TTTTCTTGGCAGCTTGGGCCATCTTGCTTAGACAGTCAGCAATTTGTCTCCTCCTTTAAAC	300					
QY	1788	TGAGTCATCATCTTAGTCCAACTAATGCAGTCGATACAATCGGTAGATAGAGAAGACCCC	1847					
Db	301	TGAGTCATCATCTTAGTCCAACTAATGCAGTCGATACAATCGGTAGATAGAGAAGACCCC	360					
QY	1848	ACGGCAGCCAGATGGGACTGGTCGTGTTTGTGCTTTCTCCAAGTCAGCACCCCAAAGT	1907					
Db	361	ACGGCAGCCAGATGGGACTGGTCGTGTTTGTGCTTTCTCCAAGTCAGCACCCCAAAGT	420					
QY	1908	CAATGCACAGACACCCCGGTGGGTGAGCGGTGCTTCTCAAACGGCCGAAGTTGCCTCT	1967					
Db	421	CAATGCACAGACACCCCGGTGGGTGAGCGGTGCTTCTCAAACGGCCGAAGTTGCCTCT	480					
QY	1968	TTTAGGAATCTCTTTGGAAATTTGGAGCAGCATGACTCTGAGTTTGAGCTATTAAAGTACT	2027					
Db	481	TTTAGGAATCTCTTTGGAAATTTGGAGCAGCATGACTCTGAGTTTGAGCTATTAAAGTACT	540					
QY	2028	TCTTACACATTGAAAAAIAAAAAAAAAA	2054					
Db	541	TCTTACAAAAAIAAAAAAAAAAIAAAAA	567					

Search completed: March 12, 2004, 06:15:07  
Job time : 484.733 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 10, 2004, 09:18:38 ; Search time 3400.62 Seconds  
(without alignments)  
18036.975 Million cell updates/sec

Title: US-10-084-817-32

Perfect score: 2054

Sequence: 1 cgggcctgcctgcgc.....cattgaaaaaaaaaaaaa 2054

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: em\_estba:\*

2: em\_esthum:\*

3: em\_estin:\*

4: em\_estmu:\*

5: em\_estov:\*

6: em\_estpl:\*

7: em\_estro:\*

8: em\_htc:\*

9: gb\_est1:\*

10: gb\_est2:\*

11: gb\_htc:\*

12: gb\_est3:\*

13: gb\_est4:\*

14: gb\_est5:\*

15: em\_estfun:\*

16: em\_estom:\*

17: em\_gss\_hum:\*

18: em\_gss\_inv:\*

19: em\_gss\_pln:\*

20: em\_gss\_vrt:\*

21: em\_gss\_fun:\*

22: em\_gss\_man:\*

23: em\_gss\_mus:\*

24: em\_gss\_pro:\*

25: em\_gss\_rod:\*

26: em\_gss\_phg:\*

27: em\_gss\_vrl:\*

28: gb\_gss1:\*

29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1703.6	82.9	1911	11	BC011352 Homo sapi
2	1173	57.1	1753	11	AK090320 Mus muscu
C 3	1051.2	51.2	1201	13	BC334954 BX334954
C 4	1048.8	51.1	1201	13	BC356240 BX356240

5	1027.4	50.0	1201	13	BC356241
6	1021.6	49.7	1146	13	BC423030
C 7	1016.8	49.5	1132	13	BC423029
C 8	1013.6	49.3	1201	13	BC381058
C 9	994	48.4	1090	13	BC423083
C 10	988.8	48.1	1201	13	BC339704
C 11	973.8	47.4	1106	13	BC423084
C 12	973	47.4	973	29	AY406047
C 13	941.4	45.8	992	13	BC401342
C 14	940.2	45.8	1201	13	BC334955
C 15	938.4	45.7	1201	13	BC380623
C 16	935.4	45.5	1201	9	AL533136
C 17	929.6	45.3	1178	13	BC381846
C 18	919	44.7	1201	13	BC417479
C 19	915.4	44.6	1001	9	AL568473
C 20	914	44.5	1201	13	BC439802
C 21	912	44.4	1201	13	BC336059
C 22	907.6	44.2	1007	9	AL540553
C 23	901.4	43.9	1201	13	BC379690
C 24	896.6	43.7	940	13	BC340279
C 25	891.4	43.4	1201	9	AL540253
C 26	889.4	43.3	1147	13	BC381845
C 27	882.4	43.0	1201	9	AL571001
C 28	875	42.6	1201	13	BC339705
C 29	869	42.3	1147	14	CK232729
C 30	867	42.2	1201	9	AL540254
C 31	866.8	42.2	1006	13	BC419376
C 32	864.8	42.1	1201	9	AL564978
C 33	864.4	42.1	950	9	AL542818
C 34	863.2	42.0	1201	13	BC381059
C 35	860.8	41.9	1201	13	BC334775
C 36	853.6	41.6	913	9	AL569639
C 37	850.6	41.4	1147	9	AL550616
C 38	840.2	40.9	902	9	AL578595
C 39	839	40.8	1026	13	BC401343
C 40	839	40.8	1201	13	BC334774
C 41	834.2	40.6	919	9	AL556304
C 42	832	40.5	974	9	AL566107
C 43	831.6	40.5	1201	13	BC364740
C 44	829.2	40.4	1201	13	BC343677
C 45	829	40.4	881	14	CDS18132

## ALIGNMENTS

RESULT 1  
BC011352  
LOCUS  
DEFINITION  
Homo sapiens protease, serine, 11 (IGF binding), mRNA (cdna clone IMAGE:4177882), containing frame-shift errors.  
ACCESSION  
BC011352  
VERSION  
BC011352.1 GI:15030191  
KEYWORDS  
HTC.  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
1 (bases 1 to 1911)  
AUTHORS

BC011352 1911 bp mRNA linear HTC 19-NOV-2003  
Homo sapiens protease, serine, 11 (IGF binding), mRNA (cdna clone IMAGE:4177882), containing frame-shift errors.  
ACCESSION  
BC011352  
VERSION  
BC011352.1 GI:15030191  
KEYWORDS  
HTC.  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
1 (bases 1 to 1911)  
AUTHORS

Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Pailey, J., Helton, E., Kettelman, M., Madan, A., Rodrigues, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,

Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,  
 Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,  
 Butterfield, Y.S., Krzyzinski, M.I., Skalska, U., Smallos, D.E.,  
 Schnerch, A., Schein, J.E., Jones, S.J., and Matra, M.A.  
 Generation and initial analysis of more than 15,000 full-length  
 human and mouse cDNA sequences  
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)  
 2388257  
 12477932  
 2 (bases 1 to 1911)  
 Strausberg, R.  
 Direct Submission  
 Submitted (25-JUL-2001) National Institutes of Health, Mammalian  
 Gene Collection (MGC), Cancer Genomics Office, National Cancer  
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
 USA  
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
 Contact: MGC help desk  
 Email: [cgapbs-re@mail.nih.gov](mailto:cgapbs-re@mail.nih.gov)  
 Tissue Procurement: David N. Louis, M.D.  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Baylor College of Medicine Human Genome  
 Sequencing Center  
 Center code: BCM-HGSC  
 Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>  
 Contact: [amg@bcm.tmc.edu](mailto:amg@bcm.tmc.edu)  
 Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Louisseged, H.,  
 Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,  
 A.N., Gibbs, R.A.  
 Clone distribution: MGC clone distribution information can be found  
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Series: IRAX Plate: 12 Row: h Column: 8  
 This clone was selected for full length sequencing because it  
 passed the following selection criteria: matched mRNA gi: 21327712  
 This clone has the following problem: frame shifted.  
 Location/Qualifiers  
 1..1911  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4177882"  
 /tissue\_type="Brain, anaplastic oligodendroglioma with  
 1p/19q loss"  
 /clone\_lib="NCI CGAP\_Brn67"  
 /lab\_host="DH10B"  
 /note="Vector: pCMV-SPORT6"

QY	597	GGTTTCATATCGAATGTTTTCGAAAGCTTCGGTTTCTTAAACGAGAGGTGCGGTGGCTAG	656
DB	444	GGTTTCATATCGAATGTTTTCGAAAGCTTCGGTTTCTTAAACGAGAGGTGCGGTGGCTAG	503
QY	657	TGGGTCTGGGTTTATTGTTGTCGGAAGATGGACTGATCGTGAACAAATGCCACCTGGTGAC	716
DB	504	TGGGTCTGGGTTTATTGTTGTCGGAAGATGGACTGATCGTGAACAAATGCCACCTGGTGAC	563
QY	717	CAACAAGACCGGTCCTCAAGTTGAGCTGAAGAACGGTGGCCTTACGAGCCCAATCAA	776
DB	564	CAACAAGACCGGTCCTCAAGTTGAGCTGAAGAACGGTGGCCTTACGAGCCCAATCAA	623
QY	777	GGATGTGATGAGAAAGCAGACATCGCACTCATCAAAATTGACCAACGAGGCAAGTGCC	836
DB	524	GGATGTGATGAGAAAGCAGACATCGCACTCATCAAAATTGACCAACGAGGCAAGTGCC	683
QY	837	TGTCCTGCTGTTGGCCGCTCTCAGAGCTGGGCGGGAGAGTTGTTGGTTCGCCATCGG	896
DB	684	TGTCCTGCTGTTGGCCGCTCTCAGAGCTGGGCGGGAGAGTTGTTGGTTCGCCATCGG	743
QY	897	AAGCCCGTTTTCCTTCAAAAACACAGTCAACACCGGGATCGTGAGCACCACCCAGGAGG	956
DB	744	AAGCCCGTTTTCCTTCAAAAACACAGTCAACACCGGGATCGTGAGCACCACCCAGGAGG	803
QY	957	CGGCAAAAGAGTGGGGTCCGGCACTCAGACATGAGCTACATCCAGACCGACCCATCAT	1016
DB	804	CGGCAAAAGAGTGGGGTCCGGCACTCAGACATGAGCTACATCCAGACCGACCCATCAT	863
QY	1017	CAACTATGAAACTCGGGAGGCCCTTAGTAAACCTGGAACGGTGAAGTGAATTA	1076
DB	864	CAACTATGAAACTCGGGAGGCCCTTAGTAAACCTGGAACGGTGAAGTGAATTA	923
QY	1077	CACTTTGAAGTGAAGTGAATCTCTTTGCAATCCATCTGATGAAGTGAAGTGAAT	1136
DB	924	CACTTTGAAGTGAAGTGAATCTCTTTGCAATCCATCTGATGAAGTGAAGTGAAT	983
QY	1137	CCTCAGGAGTCCCATGACCGACAGCCAAAGGAAAGCATCACCAGAAAGAGTATAT	1196
DB	984	CCTCAGGAGTCCCATGACCGACAGCCAAAGGAAAGCATCACCAGAAAGAGTATAT	1043
QY	1197	TGGTATCCGAATGATGTCACTCAGTCCAGGAAAGCAAGAGCTGAAGACCGGACCG	1256
DB	1044	TGGTATCCGAATGATGTCACTCAGTCCAGGAAAGCAAGAGCTGAAGACCGGACCG	1103
QY	1257	GGACTTCCAGACGCTGATCTCAGGAGCGTATATAATTAAGTAAATCTCTGATACCCAGC	1316
DB	1104	GGACTTCCAGACGCTGATCTCAGGAGCGTATATAATTAAGTAAATCTCTGATACCCAGC	1163
QY	1317	AGAAGCTGGTGTCTCAAGGAAAACGAGCTCATATCAGATCAATGACAGTCCGTGGT	1376
DB	1164	AGAAGCTGGTGTCTCAAGGAAAACGAGCTCATATCAGATCAATGACAGTCCGTGGT	1223
QY	1377	CTCCGCCAATGATGTGACGAGCGTCAATTAAGGAAAGCAACCTGAAACATGGTGGTCCG	1436
DB	1224	CTCCGCCAATGATGTGACGAGCGTCAATTAAGGAAAGCAACCTGAAACATGGTGGTCCG	1283
QY	1437	CAGGGTGAATGAAGATATCATGATCAAGTGAATCCGAAAGAAATGACCCATAGGAGGA	1496
DB	1284	CAGGGTGAATGAAGATATCATGATCAAGTGAATCCGAAAGAAATGACCCATAGGAGGA	1343
QY	1497	GGCATGAGTGGACTTCATGTTTCCCTCAAGACTCTCCCGTGGATCAGGATGAGGACT	1556
DB	1344	GGCATGAGTGGACTTCATGTTTCCCTCAAGACTCTCCCGTGGATCAGGATGAGGACT	1403
QY	1557	CTGGGCTGTGGAATAGGACACTCAAGACTTTTGAATGATGCTTTGTTGTTGTTGTTGAG	1616
DB	1404	CTGGGCTGTGGAATAGGACACTCAAGACTTTTGAATGATGCTTTGTTGTTGTTGTTGAG	1463
QY	1617	ACTTCCCTGCCCAACAGAAATCTCTTGTAGATTTGAGGCAAAACAAATGTAATGTTGCA	1676
DB	1464	ACTTCCCTGCCCAACAGAAATCTCTTGTAGATTTGAGGCAAAACAAATGTAATGTTGCA	1523





		Matches 1467;	Conservative 0;	Mismatches 275;	Indels 29;	Gaps 10;
QY	265	GGCGCCCGTCCGCTCAGAGGCGCGTCCGCGAGGCGCTGACAGTGGTGGCC				324
Db	1	GGTACAGCTGCGGCTCAGAGGCTCCCTCGCGAGGCGCTGAAATGCGTAGTGGCC				60
QY	325	TTGGGGTGGCAGGCTCGGCCACGCTCGGCGCGCGCGCGCGCGCGCTCTGTGTGC				384
Db	61	TTGGGGTGGCAGGCTCGGCCACAGTACGACGCGCGCGCACAGCGCGCGCTGTGCGTGT				120
QY	385	GCAGCAGGAGCGGCTGGCGGAGCGGAGCGGACACCTACGCGCAACCTGTGCCAGCTG				444
Db	121	GCAGCAGGAGCGGCTGGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG				180
QY	445	CGGCGCCGAGCCGCTCCGAGAGGCTGACCGGCGCCCGCTCATCTCTCTGCGAGCGC				504
Db	181	CGGCGCCGAGCCGCTCCGAGAGGCTTCGCCAGCGCCCGCTCATCTCTCTGCGAGCGC				240
QY	505	GGAGCCTGCGGCCAAGGCGAGAGATCCCAACAGTTTTCGCGCATATAATTAATCTTATC				564
Db	241	GGCGCTGCGGCCAAGGCGAGAGATCCCAACAGTTTTCGCTCATAGTACAACTTAT				300
QY	565	GGGAGCTGTGGAGAGATGCCCGCTGCGGCTGCTCATATCGAATGTTTTCGCAAGCTT				624
Db	301	GCTGATGTGGAGAGATGCCCGCTGCTGCTGCTCATATCGAATGTTTTCGCAAGCTT				360
QY	625	CCGTTTCTAAACGAGAGTGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT				684
Db	361	CTTTCTCGAAGAGGAGTGCCAGTGCCAGTGCCAGTGCCAGTGCCAGTGCCAGTGCC				420
QY	685	GGACTGATGTGACAAATGCCAGCTGGTGACCAACAGCAGCGGTCGAAAGTTGAGCTG				744
Db	421	GGACTGATGTGACAAATGCCAGCTGGTGACCAACAGCAGCGGTCGAAAGTTGAGCTG				480
QY	745	AAGAACGGTGCACCTTACGAGCCGCAAAATCAAGGATGTGGATGAGAGCAGACATCGCA				804
Db	481	AAGAACGGTGCACCTTACGAGCCGCAAAATCAAGGATGTGGATGAGAGCAGACATCG				540
QY	805	CTCATCAAAATGACCAACAGGCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT				864
Db	541	CTTATCAAGATGTGACCAAGGAGAGCTGCGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT				600
QY	865	CTGGCGCGGAGAGTTCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT				924
Db	601	CTGAGACTGAGAGATTTGTAGTTCCTATTTGAAGCCCTTTCTCTCTTCAAAACACAGTC				660
QY	925	ACCAACGGATGCTGAGCACCACAGCAGGAGCGGCGGAGAGCTGGGCTCCGCACTCA				984
Db	661	ACCACTGGATGCTGAGCACCACAGCAGGAGCGGCGGAGAGCTGGGACTTCGGAATCC				720
QY	985	GACATGGAATACATCCAGACCGACCCATCATCAACTATGGAATCTGGGAGCCCGTTA				1044
Db	721	GATATGGAATACATCCAGACCGACCCATCATCAACTATGGAATCTGGGAGCCCGTTA				780
QY	1045	GTAACCTGGAGTGAAGTATGGAATTAACATTTTCAAGAGTGAAGTGAATCTCC				1104
Db	781	GTAACCTGGAGTGAAGTATGGAATTAACATTTTCAAGAGTGAAGTGAATCTCC				840
QY	1105	TTTCAATTCCTATGATAAGATTAAAGTTTCTTACGAGGTCCTATCCAGACGCTGATCTCAGAGCG				1164
Db	841	TTTCAATTCCTATGATAAGATTAAAGTTTCTTACGAGGTCCTATCCAGACGCTGATCTCAGAGCG				900
QY	1165	AAAGGAAAGCATCACCAAGAGAGATATTTGGTATCCGATGATGTCACCTCACGCTCC				1224
Db	901	AAAGGAAAGCATCACCAAGAGAGATATTTGGTATCCGATGATGTCACCTCACATCT				960
QY	1225	AGCAAGCAAGAGCTGAGAGCCGCGACCGGAGCTTCCAGACGCTGATCTCAGAGCG				1284
Db	961	AGCAAGCAAGAGCTGAGAGCCGCGACCGGAGCTTCCAGACGCTGATCTCAGAGCG				1020
QY	1285	TATATAATTGAATATCTCTGATACCCAGAGAGCTGGTGGTCTCAAGGAAACGAC				1344
Db	1021	TATATAATTGAATATCTCTGATACCCAGAGAGCTGGTGGTCTCAAGGAAATGAC				1080

QY	1345	GTCTAATACATCAATGAGCAGTCCGTGGTCTCCGCAATGATGTACGAGCAGTCA	1404
Db	1081	GTCTAATACATCAATGAGCAGTCTGTGTCTCTGCTCACTGCCAATGACCTCAGGATGTATC	1140
QY	1405	AAAAGGAAAGCAGCTGAAATGATGTGTCCGAGGAGTAAATGAAGATATCATGATCACA	1464
Db	1141	AAAAGGAGAAACAGCTGAAATGATGTGTCCGAGGAGTAAATGAAGATATGTGATTAAC	1200
QY	1465	GTGATTTCCGAAAGAAATGACCCATAGGCAGAGGATGAGCTGGAGCTTTCATGTTTCCCTC	1524
Db	1201	GTGATTTCCGAAAGAAATGACCCATAGGCAGAGGATGAGCTGGAGCTTTCATGTTTCCCTC	1260
QY	1525	AAAGACTCTCCCGTGGATGACGAGTCTGAGTCTGGCTGCTGGAATAGACACTCAAGA	1584
Db	1261	AAAGA--CTCAGGAGATGGCGCATGAGAA--CTGAGCTGGTGACGAGGACACCCGGGA	1316
QY	1585	CTTTTGATGTCATTTTGTGTTTTCAGTGGAGACTCCCTTGGGCAACAGAAATCTTCTTGA	1644
Db	1317	CTTTTGTCACAACTTTTGTGTTTCA--GGGAACACCCCTGCGCAGAGATCTTCTTGA	1374
QY	1645	TAGTTTGCAGCAAAACAAATGTAATGTTGACATCTTACGAGTCCGAGGAGCTTCCCTTCT	1704
Db	1375	TAGTTTGCAGCAAAACAAATGTAATGTTGACGAGTCCGAGGAGCTTCCCTTCT	1433
QY	1705	GTATCTTATGATGACGTGTGCTTTTCTTCCAGCTTCCGCTTCCCTTACAGT	1764
Db	1434	GTATCTTATGATGACGTGTGCTTTTCTTCAAGCTTGGGCTGTCTCTTACAGT	1493
QY	1765	CAGATTTGCTCTCTCTTAACTGACATCATCTTACGAGTCCGAGTCCGAGTAC	1824
Db	1494	CAGATTTGCTCTCTCTTAACTGACATCATCTTACGAGTCCGAGTCCGAGTAC	1545
QY	1825	AATCGTAGATAGAAAGCCCGGAGCAGGAGTGGAGTGGCTGTGCTGCTT	1884
Db	1546	ATCGTAGGCAAGAGGAGGCTCTC---GCAGCCTATGCTTGGGCTGTGCTTACACTTT	1601
QY	1885	T-CTCAGATCAGACACCCAAAGG---TCAATGACAGAGACCCCGGCTGGGCTG	1940
Db	1602	TCCTCCGAGTGGACTCAGAGGATTTGATGCCCGGAGACACAGTGGGTGA-CCCTG	1660
QY	1941	GCTTCTCAACGGCGAAGTGGCTCTTTTGAAGATCTCTTTGGAATGGGAGCAGATG	2000
Db	1661	GCTTCTGAGATGGGCAAAATGCTCTT---AGGAATCTCTGGAACCCGAGTACAGT	1715
QY	2001	ACTCTGAGTTTGGCTATTAAAGTACTTCTT	2031
Db	1716	ACTTGTGATTTGAGCTATTAAATACTTCTT	1746

RESULT 3  
BX334954/c  
LOCUS BX334954 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA  
DEFINITION clone CS01009YE13 3-PRIME, mRNA sequence.  
ACCESSION BX334954  
VERSION BX334954.1 GI:30343366  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 1201)  
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished (2001)  
COMMENT Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: segre@genoscope.cns.fr, Web: www.genoscope.cns.fr  
Library was constructed by Life Technologies, a division of  
Invitrogen. This sequence belongs to sequence cluster 5631.f For  
more information about this cluster, see

<http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CSODI009AC07NP1&cluster=5631.f>. Contact : Feng Liang Email : fliang@lifetech.com URL : <http://fulllength.invitrogen.com/invidrogen Corporation 1600 Faraday Avenue Genoscope sequence ID : CSODI009AC07NP1>

FEATURES	Location/Qualifiers
source	1..1201 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="CSODI009YE13" /tissue_type="PLACENTA COT 25-NORMALIZED" /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED" /note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match	Score	DB 13	Length	Matches	Conservative	Mismatches	Gaps	Indels
Best Local Similarity	51.2%	1051.2	1201	88.7%	5.6e-160	60	2	1;
Matches 1065;						74		
QY	805	CTCATCAAAATTGACACACAGGCGAAGCTGGCTGTCTGCTTGCGGGCCTCCAGAG	864	:	:	:	:	:
Db	1200	MTCRWYAATTTTRACAACAAGG--CAAGYKCTGCACCTGTGYTTGGCGSCCCC	1143	:	:	:	:	:
QY	865	CTGGCGCGGGGAGAGTTCGTGGTCGCCATCGGAAGCCCGTTTTCCCTTCAAACAACAGTC	924	:	:	:	:	:
Db	1142	CGCGSGCGGGGAGAKTCKKGKYCGCCAAGGAAACCVYKTTCCTTWCAAAAAACAKTC	1083	:	:	:	:	:
QY	925	ACCACCGGATCGTGAGCACCAACCACGAGGCGGAAGAAGCTGGGCTCGCAACTCA	984	:	:	:	:	:
Db	1082	ACCACCGGTATYKTGAGCACCAACCACGAGGCGGMAAAGAGYTGGGSYCCGCAATTC	1023	:	:	:	:	:
QY	985	GACATGGACTACATCCAGACCGAGCCATCATCACTATGGAAGCTCGGAGGCGCGTTA	1044	:	:	:	:	:
Db	1022	GACATGGACTACATCCAGACCGAGCCATCATCACTATGGAAGCTCGGAGGCGCGTTA	963	:	:	:	:	:
QY	1045	GTAAACCTGGACGGTGAAGTGATTGGAATTAACACTTTGAAAGTGACAGCTGGAATCTCC	1104	:	:	:	:	:
Db	962	GTAAACCTGGACGGTGAAGTGATTGGAATTAACAYTTTGAAAGTGACAGCTGGAATCTCY	903	:	:	:	:	:
QY	1105	TTTGCAATCCCATCTGATAGATTAAAAGTTCTCAGGAGTCCCATGACCGCAGAGGCC	1164	:	:	:	:	:
Db	902	TYKGCAATCCCATCTGATAGATTAAAAGTTCTCAGGAGTCCCATGACCGCAGAGGCC	843	:	:	:	:	:
QY	1165	AAAGGAAAGCCATCAACCAAGAGAAGTATATTGGTATCCGAATGATGTCTACAGTCC	1224	:	:	:	:	:
Db	842	AAAGGAAAGCCATCAACCAAGAGAAGTATATTGGTATCCGAATGATGTAACTACAGTCC	783	:	:	:	:	:
QY	1225	AGCAAGGCCAAAGAGCTGGAAGGACCGGCACCGGACTTTCCACAGAGCTGATCTCAGGAGG	1284	:	:	:	:	:
Db	782	AGMAAGGCCAAAGAGYTGGAAGACCGGCACCGGATTTTACCAGACKTGATTTTCAGGAGCG	723	:	:	:	:	:
QY	1285	TATATAATTGAAGTATTCCTGATACCCAGCAGAGTGCTGGTCTCAAGAAACGAC	1344	:	:	:	:	:
Db	722	TATATAWTGAAGTAWTTCVTGATACCCAGCAAKYTTGGTKGTGTCAAGGAAACGAC	663	:	:	:	:	:
QY	1345	GTCAATAATCAGCATCAATGACAGTCCTGGTCTCCGCCAATGATGTCTCAGCGACGTCATT	1404	:	:	:	:	:
Db	662	GTWTATWTGAGCATCAATGACAGTCCTGGTCTCCGCCAATGATGTCTCAGCGACGTSATG	603	:	:	:	:	:
QY	1405	AAAAGGGAAGCACCTGAACTGGTGTCCGACAGGGGTAAATGAAGATATCATGATCACA	1464	:	:	:	:	:
Db	602	AAAAGGGAAGCACCCNGRAACANGGNGCCGACAGGGGNAANGAAGATANCAGGANGACA	543	:	:	:	:	:
QY	1465	GTGATTCGGAAGAAATTGACCCATAGGCAGAGCATGAGCTGGACTTCATGTTTCCCCCTC	1524	:	:	:	:	:
Db	542	GTGATTCGGAAGAAATTGACCCATAGGCAGAGCATGAGCTGGACTTCATGTTTCCCCCTC	483	:	:	:	:	:
QY	1525	AAAGACTCTCCCGTGATGACGGATGAGGACTCTGGGCTGTGGTAATAGCACTCAAGA	1584	:	:	:	:	:

Db		482	A A G A G C C C C C G T G A T G A C G G A T G A G A C C C T G G C G C G G A T A G A S A C C C A A G A	423
Qy		1585	C T T T T G A C T G C C A T T T T G T T G T T C A G T G G A G A C T C C C T G G C C A A C A G A A N C T T C T T T G A	1644
Db		422	C T T T T G A C T G C C A T T T T G T T T G G C C A G T G G A G A C T C C C T G G C C A A C A G A A T C C T T C T T T G A	363
Qy		1645	T A G T T T T G C A G G C A A A A C A A A T G T A A T G T T G C A G A T C C G C A G G C A G A A G C T C T G C C C T T C T	1704
Db		362	T A G T T G C G A G G C A A A A A A A A T G T A A T G T G C A G A T C C G C A G G C A G A A G S T C T G C C K C T C T	303
Qy		1705	G T A T C C T A T G T A T G C A G T G C T T T T T C T T G C C A G C T T G G G C C A T T C T T G C T T T A G A C A G T	1764
Db		302	G T A T C C T A T G T A A G C A G T G G G C T T T G T T G C C A G C T T G G G C C A R C T G S C T A A G A C A G T	243
Qy		1765	C A G C A T T T G T C C T C C T T T A A C T G A G T C A T C A T C T T A G T C C A A C T A A T C C A G T C G A T A C	1824
Db		242	C A G C A T T T G T C T C T C C T T T A A C T G A G T C A T C A T C T T A G T C C A A C T A A G C A G C T G C A T A C	183
Qy		1825	A A T G C C T A G A T A G A A A A G C C C C A C G G G A G C C A G G A T G G A C T T G G T C G T G T T T G T G C T T T	1884
Db		182	A A G G C G T A G A T A G A A A A G C C C C A C G G G A C C A G G A T G G A C T G G T C G T G T T T G T G C T T T	123
Qy		1885	T C T C C A A G T C A G C A C C A A A G G T C A A T G C A C A G A G A C C C G G G T G G G T A G C C T G G C T T	1944
Db		122	T C T C C A A G T C A G C A C C A A A G G T C A A T G C A C A G A G A C C C G G G T G G G T A G C C T G G C T T	63
Qy		1945	C T C A A A C G G C C G A A G T T G C C T C T T T T A G A A T C T T T T G G A A T C T T T T G G A A T T G G A G C A C A T A G A C T	2004
Db		62	S T C A A A C G G C G A A G T T G C G T C T T T T A G A A T C T T T T G G A A T C T T T T G G A A T T G G A G C A S G A T K A G T G	3
Qy		2005	T	2005
Db			T	T
RESULT 4				
EX356240/c				
LOCUS				
DEFINITION			1201 bp mRNA linear EST 05-MAY-2003	
ACCESSION			BX356240 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA	
VERSION			clone CSOD1009YL10 3-PRIME, mRNA sequence.	
KEYWORDS			EST.	
SOURCE			GI:30380054	
ORGANISM			Homo sapiens (human)	
REFERENCE			Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
AUTHORS			Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
TITLE			L1.(bases 1 to 1201)	
JOURNAL			I.L.W.B., Gruber,C., Jessee,J. and Polayes,D.	
COMMENT			Full-length cDNA libraries and normalization Unpublished (2001) Contact: Genoscope Genoscope - Centre National de Sequencage BP 191 91006 EVRY cedex - France Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr Library was constructed by life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 5631.f For more information about this cluster, see <a href="http://www.genoscope.cns.fr/">http://www.genoscope.cns.fr/</a> <a href="http://cluster.cgi?seq=CSOD1009DF05NP1&amp;cluster=5631.f">http://cluster.cgi?seq=CSOD1009DF05NP1&amp;cluster=5631.f</a> . Contact : Feng Liang Email : fliang@lifetech.com URL : <a href="http://fulllength.invitrogen.com/">http://fulllength.invitrogen.com/</a> Invitrogen Corporation 1600 Faraday Avenue Genoscope sequence ID : CSOD1009DF05NP1. Location/Qualifiers 1..1201 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="CSOD1009YL10" /tissue_type="PLACENTA COT 25-NORMALIZED" /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED" /notes="1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime end enriched, double-strand cDNA was	
FEATURES				
source				









```

Db 345 TGAATGTTGCAGATCCGAGGAGAGCTCTCCCTTCTGTATCCTATGATGAGGT 286
Qy 1725 GCTTTTCTTGCCAGCTTGGCCATTCTTCTTAGACAGTCAAGATTTGTCTCTCTTT 1784
Db 285 GCTTTTCTTGCCAGCTTGGCCATTCTTCTTAGACAGTCAAGATTTGTCTCTCTTT 226
Qy 1785 AACTGAGTCATCTTAGTCCAACTAAATGACGTCGATCAATCGGTAGTAGAGAGC 1844
Db 225 AACTGAGTCATCTTAGTCCAACTAAATGACGTCGATCAATCGGTAGTAGAGAGC 166
Qy 1845 CCCAGGGAGCAGGATGGGACTGGTGTGTTTGTCTTCTCAAGTCAGCACCCAAA 1904
Db 165 CCCAGGGAGCAGGATGGGACTGGTGTGTTTGTCTTCTCAAGTCAGCACCCAAA 106
Qy 1905 GGTCAATGACAGAGACCCCGGTGGTGGAGCGCTGCTCTCAAGCGCGAGTGGC 1964
Db 105 GGTCAATGACAGAGACCCCGGTGGTGGAGCGCTGCTCTCAAGCGCGAGTGGC 46
Qy 1965 TCTTTTAGGAATCTCTTTGGAAATGGGAGCAGATGACTCTGAGT 2009
Db 45 TCTTTTAGGAATCNCNTNGGAANNNGGAGCAGCAGAGAGACTKMK 1

RESULT 8
BX381058/c 1201 bp mRNA linear EST 08-MAY-2003
LOCUS BX381058 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
DEFINITION clone CS0D1060YC24 3-PRIME, mRNA sequence.
ACCESSION BX381058
VERSION BX381058.1 GI:30456926
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1201)
AUTHORS Li, W.B., Gruber, C., Jesssee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 5631.f For
more information about this cluster, see
http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0D1060B12NP1&cluster=5631.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0D1060B12NP1.

FEATURES
Location/Qualifiers
1..1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1060YC24"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with NotI and cloned into the NotI and EcoRV
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
Query Match 49.3%; Score 1013.6; DB 13; Length 1201;
Best Local Similarity 96.3%; Pred. No. 6.7e-154;
Matches 1054; Conservative 10; Mismatches 27; Indels 4; Gaps 3;

Qy 911 TTCAAAACACAGTCACCCCGGATCTGTAGACACACCCAGCGCGGCAAGAGCTGG 970
Db 1106 TTCCTTTAAAACACATCCGCCGGATCGTGACACACCCAGCGCGGCAAGAGCTGG 1047

```

```

Qy 971 GGCTCCGAACTCAGACATGGAATCATCATCCAGACCGAGCCCATCATCACTATGGAAC 1030
Db 1046 GS--CCGCAACTCAGAYWT-GACTACATCCAGACCGAGCCGCTCATCACTATGGAAC 990
Qy 1031 CGGGAGCGCGGTAGTAAACCTCGGAGCGGTGAAGTGAATGGAATTAACACTTTTGAAGTGA 1090
Db 989 CGGGAGCGCGGTAGTAAACCTCGGAGCGGTGAAGTGAATGGAATTAACACTTTTGAAGTGA 930
Qy 1091 CAGCTGGAATCTCTTTGCAATCCATCTGATTAAGATTAAGAGTTTCTTCAAGAGTCCC 1150
Db 929 CAGCTGGAATCTCTTTGCAATCCATCTGATTAAGATTAAGAGTTTCTTCAAGAGTCCC 870
Qy 1151 ATGACCGACAGGCGCAAGGAAAGCCATCACCAAGAAAGATATATTTGGTATCCCAATGA 1210
Db 869 ATGACCGACAGGCGCAAGGAAAGCCATCACCAAGAAAGATATATTTGGTATCCCAATGA 810
Qy 1211 TGTCACTCAGCTCCAGCAAGCCAAAGAGCTGAAGAGCCGCGACCCGGAGACTTCCAGAGC 1270
Db 809 TGTCACTCAGCTCCAGCAAGCCAAAGAGCTGAAGAGCCGCGACCCGGAGACTTCCAGAGC 750
Qy 1271 TGATCTCAGGAGCGTATATTAATTGAAGTAAATCTCTGATACCCAGCAGAAAGTGGTGGTC 1330
Db 749 TGATCTCAGGAGCGTATATTAATTGAAGTAAATCTCTGATACCCAGCAGAAAGTGGTGGTC 690
Qy 1331 TCAAGAAACAGCGTCAATTAATCAGCATCAATGAGCAGTCCGTGGTCTCCGCCAATGATG 1390
Db 689 TCAAGAAACAGCGTCAATTAATCAGCATCAATGAGCAGTCCGTGGTCTCCGCCAATGATG 630
Qy 1391 TCAGCGACGTCATTAAAGGGGAAAGCACCCTGAAATGCTGGTCCGAGGGGTAAATGAAG 1450
Db 629 TCAGCGACGTCATTAAAGGGGAAAGCACCCTGAAATGCTGGTCCGAGGGGTAAATGAAG 570
Qy 1451 ATATCATGATCAAGTGAATTCGGAAGAAATGACCATAGGAGAGGATGAGCTGGAC 1510
Db 569 ATATCATGATCAAGTGAATTCGGAAGAAATGACCATAGGAGAGGATGAGCTGGAC 510
Qy 1511 TTCATGTTTCCCTCAAAGACTCTCCCGTGGATGACGAGTGAAGACTCTGGGTCTGTGGAA 1570
Db 509 TTCATGTTTCCCTCAAAGACTCTCCCGTGGATGACGAGTGAAGACTCTGGGTCTGTGGAA 450
Qy 1571 TAGGACACTCAAGACTTTTGAATGCTCCATTTGTTTGTTCAGTGGAGACTCTCTGGCCAC 1630
Db 449 TAGGACACTCAAGACTTTTGAATGCTCCATTTGTTTGTTCAGTGGAGACTCTCTGGCCAC 390
Qy 1631 AGAATCTTCTTCAATGATTTGAGGAGCAAAACAAATGTAATGTTGAGATCCCGAGGACAG 1690
Db 389 AGAATCTTCTTCAATGATTTGAGGAGCAAAACAAATGTAATGTTGAGATCCCGAGGACAG 330
Qy 1691 AGCTCTGCCCTTCTGATCTCTATGATGAGTGTGCTTTTCTTCCAGACTTTGGGCAAT 1750
Db 329 AGCTCTGCCCTTCTGATCTCTATGATGAGTGTGCTTTTCTTCCAGACTTTGGGCAAT 270
Qy 1751 CTTGCTTAGACAGTCAAGCTTTGCTCTCTCTTAACTGAGTCAATCATCTTAGTCCCACT 1810
Db 269 CTTGCTTAGACAGTCAAGCTTTGCTCTCTCTTAACTGAGTCAATCATCTTAGTCCCACT 210
Qy 1811 AATGCACTCGATCAATCGGTAGATAGAAGAGCCCAAGGAGCAGGATGGAGTGGT 1870
Db 209 AATGCACTCGATCAATCGGTAGATAGAAGAGCCCAAGGAGCAGGATGGAGTGGT 150
Qy 1871 CGTGTGTTGCTTTTCTTCAAGTCAGCACCCAAAGGTCATGTCAGACAGACCCCGGCTGG 1930
Db 149 CGTGTGTTGCTTTTCTTCAAGTCAGCACCCAAAGGTCATGTCAGACAGACCCCGGCTGG 90
Qy 1931 GTGAGCGCTGGCTTCTCAAAAGCGCGGAGTGGCTCTTTTAGGAATCTCT-TTGAATTTG 1989
Db 89 GTGAGCGCTGGCTTCTCAAAAGCGCGGAGTGGCTCTTTTAGGAATCTCTTNNTAGAATTTG 30
Qy 1990 GGAGCAGATGACTC 2004
Db 29 GGAGCAGATGACTC 15

```

```

RESULT 9
BX423083/c
LOCUS
DEFINITION BX423083 Homo sapiens ADULT BRAIN Homo sapiens cDNA clone
CS0DN004YC10 3-PRIME, mRNA sequence.
ACCESSION BX423083
VERSION
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE
AUTHORS 1 (bases 1 to 1090)
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 5631.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DN004BB05NP1&cluster=5631.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paradise Avenue Genoscope sequence ID : CS0DN004BB05NP1.
FEATURES
Location/Qualifiers
1..1090
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DN004YC10"
/tissue_type="ADULT BRAIN"
/dev_stage="adult"
/clone_lib="Homo sapiens ADULT BRAIN"
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-strand cDNA was digested with NotI and
cloned into the NotI and EcoRV sites of the pCMVSPORT 6
vector. Library was not normalized."
ORIGIN
Query Match 48.4%; Score 994; DB 13; Length 1090;
Best Local Similarity 98.8%; Pred.No.9.9e-151;
Matches 1019; Conservative 2; Mismatches 8; Indels 2; Gaps 2;
QY 944 CCACCCAGCGAGCGGCAAGAGCTGGGGCTCCGCAACTCAGACATGGACTCATCCAGA 1003
DB :
1074 MCACCCAGCGAGGC-GCAAAGAGCTGGGCTCCGCAAYCTCAGACAT-GACTACATCCAGA 1017
QY 1004 CCGACGCCATCATCAACTATGAAATCGGAGGCCGCTAGTAAACCTGGACGGTGAAG 1063
DB CCGACGCCATCATCAACTATGAAATCGGAGGCCGCTAGTAAACCTGGACGGTGAAG 957
QY 1064 TGAATTGGAATTAACACTTTGAAAGTGACAGCTGGAATCTCTTTGCAATCCCATCTGATA 1123
DB TGAATTGGAATTAACACTTTGAAAGTGACAGCTGGAATCTCTTTGCAATCCCATCTGATA 897
QY 1124 AGATTAAGTCTCTCAGGAGTCCCATGCGACGACGCCCAAGGAAAGCCATCACCA 1183
DB AGATTAAGTCTCTCAGGAGTCCCATGCGACGACGCCCAAGGAAAGCCATCACCA 837
QY 1184 AGAAGAGTATATTGGTATCCGAATGATGTCACTACGTCCTCCAGCAAGCCAAAGAGCTGA 1243
DB AGAAGAGTATATTGGTATCCGAATGATGTCACTACGTCCTCCAGCAAGCCAAAGAGCTGA 777
QY 1244 AGGACCGGACCGGACACTCCAGAGCTATCTCAGGAGGTATATAATTGAAGTAATTC 1303
DB AGGACCGGACCGGACACTCCAGAGCTATCTCAGGAGGTATATAATTGAAGTAATTC 717

```

```

QY 1304 CTGATACCCAGCAGAGCTGGTGTCTCAAGGAAAAACGACGTCATAATCAGCATCAATG 1363
DB :
716 CTGATACCCAGCAGAGCTGGTGTCTCAAGGAAAAACGACGTCATAATCAGCATCAATG 657
QY 1364 GACAGTCGGTGGTCTCCGCCAATGATGTCAGGAGCTGATTAAGGGAAGAACCCCTGA 1423
DB GACAGTCGGTGGTCTCCGCCAATGATGTCAGGAGCTGATTAAGGGAAGAACCCCTGA 597
QY 1424 ACATGGTGGTCCGACGAGGGTAAATGAAGATATCATGATCACAGTGATCCCGAAGAAATG 1483
DB ACATGGTGGTCCGACGAGGGTAAATGAAGATATCATGATCACAGTGATCCCGAAGAAATG 537
QY 1484 ACCCATAGGAGAGGATGAGCTGAGCTTCATGTTTCCCTCAAAGACTCTCCCGTGGATG 1543
DB ACCCATAGGAGAGGATGAGCTGAGCTTCATGTTTCCCTCAAAGACTCTCCCGTGGATG 477
QY 1544 ACGGATCAGGACTCTGGGCTCTGGAAATAGGACACTCAAGACTTTTGAAGCTGCCATTTGT 1603
DB ACGGATCAGGACTCTGGGCTCTGGAAATAGGACACTCAAGACTTTTGAAGCTGCCATTTGT 417
QY 1604 TTGTTCAAGTGGAGACTCCCTGGCCAAACAGAAATCCCTTTGTATGATGTTGAGGCAAAACAA 1663
DB TTGTTCAAGTGGAGACTCCCTGGCCAAACAGAAATCCCTTTGTATGATGTTGAGGCAAAACAA 357
QY 1664 ATGTAATGTTGCAGATCCGACGAGGAGAGCTCTGCCCTCTGTATCTATGATGACAGTG 1723
DB :
356 ATGTAATGTTGCAGATCCGACGAGGAGAGCTCTGCCCTCTGTATCTATGATGACAGTG 297
QY 1724 TGCTTTTCTTGGCAGCTTGGGCCAATCTTGTTTAGACAGTCAAGACTTTGTCTCTCTCT 1783
DB TGCTTTTCTTGGCAGCTTGGGCCAATCTTGTTTAGACAGTCAAGACTTTGTCTCTCTCT 237
QY 1784 TAACCTGAGTCATCATCTTAGTCCAACTAATGACATGATGATGATGATGATGATGATGATG 1843
DB TAACCTGAGTCATCATCTTAGTCCAACTAATGACATGATGATGATGATGATGATGATGATG 177
QY 1844 CCCACGCGAGCCAGGATGGAGCTGCTGTTTGTGTTTTCTCAAGTCAGACCCCA 1903
DB CCCACGCGAGCCAGGATGGAGCTGCTGTTTGTGTTTTCTCAAGTCAGACCCCA 117
QY 1904 AGGTCAATGACAGAGACCCCGGGTGGTGAGCGTGGCTCTCAAGCGCCGAAGTGC 1963
DB AGGTCAATGACAGAGACCCCGGGTGGTGAGCGTGGCTCTCAAGCGCCGAAGTGC 57
QY 1964 CTCTTTTAGGA 1974
DB :
56 STCTTNTAGGA 46
RESULT 10
BX339704/c
LOCUS
DEFINITION BX339704 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CS0DI075YN06 3-PRIME, mRNA sequence.
ACCESSION BX339704
VERSION BX339704.1 GI:30341839
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1201)
Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 5631.f For
more information about this cluster, see
http://www.genoscope.cns.fr/

```











Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
Library was constructed by Life Technologies, a division of  
Invitrogen. This sequence belongs to sequence cluster 5631.f For  
more information about this cluster, see  
http://www.genoscope.cns.fr/  
cgi-bin/cluster.cgi?seq=CS0DI055BB08NP1&cluster=5631.f. Contact :  
Peng Liang Email : fliang@lifetech.com URL :  
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600  
Paradise Avenue Genoscope, sequence ID : CS0DI055BB08NP1.  
Location/Qualifiers  
1. 1201  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0DI055VC16"  
/tissue\_type="PLACENTA"  
/clone\_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"  
/notes="1st strand cDNA was primed with a NotI-oligo (dT)  
primer. Five prime end enriched, double-strand cDNA was  
digested with Not I and cloned into the Not I and EcoR V  
sites of the pCMVSPORT 6 vector. Library was normalized."

FEATURES  
source

ORIGIN

Query Match 45.78; Score 938.4; DB 13; Length 1201;  
Best Local Similarity 93.78; Pred. No. 9.8e-142; Indels 3; Gaps 3;  
Matches 981; Conservative 23; Mismatches 40;

QY 981 CTCAGACATGGACTACATCAGACGCGCCATCATCAACTATGGAATCGGGAGGCC 1040  
DB 1045 CTCGCGACATKACATGACATCCAGACGCGCTCMTCACTATGAATCGGAGGCC 986  
QY 1041 GTTAGTAACCTGGACGGTGAAGTATGGAATTAACATTTGAAGTGACAGCTGGAAT 1100  
DB 985 GTTAGT-AACTGACGGTGAAGTATGGAATTAACATTTGAAGTGACAGCTGGAAT 928  
QY 1101 CTCCTTTGCAATCCCATCTGATAGATTAAGTTCTCTCAGGAGTCCCATGACCGACA 1160  
DB 927 CTCCTTTGCAATCCCATCTGATAGATTAAGTTCTCTCAGGAGTCCCATGACCGACA 868  
QY 1161 GGCCAAAGGAAGCCATCACCAGAGAGATATTTGGTATCCGAATGTCATCTCAC 1220  
DB 867 GGCCAAAGGAAGCCATCACCAGAGAGATATTTGGTATCCGAATGTCATCTCAC 808  
QY 1221 GTCCAGCAAAAGCCAAAGAGCTGAAGAGCGCGCACCGGAGCTTCCAGAGCTGATCTCAGG 1280  
DB 807 GTCCAGCAAAAGCCAAAGAGCTGAAGAGCGCGCACCGGAGCTTCCAGAGCTGATCTCAGG 748  
QY 1281 AGCGTATATTAATTGAAGTATTTCTGTATACCCAGCAGAGCTGGTCTCAGGAAA 1340  
DB 747 AGCGTATATTAATTGAAGTATTTCTGTATACCCAGCAGAGCTGGTCTCAGGAAA 688  
QY 1341 CGACGTCAATATCAGCATCAATGGACAGTCCGTGGTCTCCGCCAATGATGTCAGGACGT 1400  
DB 687 CGACGTCAATATCAGCATCAATGGACAGTCCGTGGTCTCCGCCAATGATGTCAGGACGT 628  
QY 1401 CATTAAAGGGAAGCAACCTGACATGGTGGTCCGAGGGGTATGAGATATCATGAT 1460  
DB 627 CATTAAAGGGAAGCAACCTGACATGGTGGTCCGAGGGGTATGAGATATCATGAT 568  
QY 1461 CACAGTGATTTCCCGAAGAAATTGACCCATAGGCGAGAGGATGAGCTGGACCTTCATGTTTC 1520  
DB 567 CACAGTGATTTCCCGAAGAAATTGACCCATAGGCGAGAGGATGAGCTGGACCTTCATGTTTC 508  
QY 1521 CTTCAAGAGCTCTCCCGTGGATGAGGATGAGGATCTGGGCTGGTGAATAGGACATC 1580  
DB 507 CTTCAAGAGCTCTCCCGTGGATGAGGATGAGGATCTGGGCTGGTGAATAGGACATC 448  
QY 1581 AAGACTTTTGAATGCCATTTTGTGTTTGTTCAGTGGAGACTCCCTGGCCCAACAGAAATCCTTC 1640  
DB 447 AAGACTTTTGAATGCCATTTTGTGTTTGTTCAGTGGAGACTCCCTGGCCCAACAGAAATCCTTC 388  
QY 1641 TTGATAGTTTGCAGGCAAAACAAATGTAATGTTGCAGATCCCGAGGACAGAGCTCTGCC 1700

DB 387 TTGATAGTTTGCAGGCAAAACAAATGTAATGTTGCAGATCCCGAGGACAGCTCTGCC 328  
QY 1701 TTCTGTATCCTATGATGATGAGTGTGCTTTTCTTCCAGCTTTGGCCATTTCTTGCTTAGA 1760  
DB 327 TTCTGTATCCTATGATGATGAGTGTGCTTTTCTTCCAGCTTTGGCCATTTCTTGCTTAGA 268  
QY 1761 CAGTCAGCATTTGTCTCTCTCTTAACTGAGTCATCATCTTATGTCACAACTAATGCAAGTCG 1820  
DB 267 CAGTCAGCATTTGTCTCTCTCTTAACTGAGTCATCATCTTATGTCACAACTAATGCAAGTCG 208  
QY 1821 ATACAATGCGTAGATAGAGAAGCCCAACGGGAGCCAGATGGGACTGCTGTTGTTGTG 1880  
DB 207 ATACAATGCGTAGATAGAGAAGCCCAACGGGAGCCAGATGGGACTGCTGTTGTTGTG 148  
QY 1881 CTTTCTTCCAAAGTCAGCACCCCAAGGTCAATGCACAGAGACCCCGGTCGGTGAGCGCTG 1940  
DB 147 CTTTCTTCCAAAGTCAGCACCCCAAGGTCAATGCACAGAGACCCCGGTCGGTGAGSGCTG 88  
QY 1941 GCTTCTCAAAAGCGCGAAGTTGCTCTTTTATGGAATCTCTTTGGAATTTGGGAGCAG-AT 1999  
DB 87 GCTTCTCAAAAGCGCGAAGTTGCTCTTTTATGGAATCTCTTTSSMCCCGGAGCAGNAT 28  
QY 2000 GACTCTGAGTTTGAAGTATTAAGTAC 2026  
DB 27 GCTCTGAGTTGNYGCEMYMANWAM 1

Search completed: March 11, 2004, 14:23:53  
Job time : 3406.84 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 10, 2004, 08:38:17 ; Search time 2825.63 Seconds  
(without alignments)  
17011.249 Million cell updates/sec

Title: US-10-084-817-37  
Perfect score: 1109  
Sequence: 1 cctgcacggagacactgg.....acanaattccaccctatccg 1109

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl.\*

- 1: gb.ba.\*
- 2: gb.htg.\*
- 3: gb.in.\*
- 4: gb.in.\*
- 5: gb.ov.\*
- 6: gb.pat.\*
- 7: gb.ph.\*
- 8: gb.pl.\*
- 9: gb.pr.\*
- 10: gb.ro.\*
- 11: gb.sts.\*
- 12: gb.sy.\*
- 13: gb.un.\*
- 14: gb.vi.\*
- 15: em.ba.\*
- 16: em.fun.\*
- 17: em.hum.\*
- 18: em.in.\*
- 19: em.mu.\*
- 20: em.cm.\*
- 21: em.or.\*
- 22: em.ov.\*
- 23: em.pat.\*
- 24: em.ph.\*
- 25: em.pl.\*
- 26: em.ro.\*
- 27: em.sts.\*
- 28: em.un.\*
- 29: em.vi.\*
- 30: em.htg.hum.\*
- 31: em.htg.inv.\*
- 32: em.htg.other.\*
- 33: em.htg.mus.\*
- 34: em.htg.pln.\*
- 35: em.htg.rod.\*
- 36: em.htg.mam.\*
- 37: em.htg.vrt.\*
- 38: em.sy.\*
- 39: em.htgo.hum.\*
- 40: em.htgo.mus.\*
- 41: em.htgo.other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	802	72.3	851	6	AX014337 Sequence
2	802	72.3	851	6	BD222216 Human nuc
3	699.8	63.1	1095	6	AX281854 Sequence
4	696.2	62.8	842	6	AX397528 Sequence
5	696.2	62.8	853	6	AX302539 Sequence
6	696.2	62.8	853	6	AX409461 Sequence
7	696.2	62.8	853	9	HUM927A
8	664.6	59.9	683	9	BC000897 Homo sapi
9	624.4	56.3	647	9	HSR9A927
10	378	34.1	378	9	BT007173
11	376.4	33.9	378	12	BT007946
12	362	32.6	156819	2	AP006283
13	362	32.6	175416	2	AP006286
14	362	32.6	187160	9	AC138230
15	360.4	32.5	143779	9	AC136475
16	351	31.7	352	6	AR273461
17	351	31.7	352	6	AR277042
18	351	31.7	352	6	AR407317
19	351	31.7	352	6	AX368494
20	340.8	30.7	622	4	AF272042
21	335.8	30.3	764	6	AR148104
22	335.8	30.3	764	6	AR350926
23	335.8	30.3	764	6	AR433302
24	335.8	30.3	764	6	AX316946
25	335.8	30.3	764	6	BD084479
26	302.6	27.3	695	6	AX281855
27	296.4	26.7	355	9	HSIFNIN2
28	289.6	26.1	637	9	BC006794
29	286	25.8	316	6	AX397529
30	285.4	25.7	122351	9	AC023157
31	282.4	25.5	621	6	BD190298
32	281.8	25.4	808	6	AR380734
33	281.8	25.4	808	6	BD190297
34	281.8	25.4	808	9	HS18U
35	280.2	25.3	222085	2	AC144988
36	278.6	25.1	749	9	BC008417
37	277.8	25.0	645	9	BC022439
38	277.2	25.0	905	6	BD190334
39	275.6	24.9	905	6	BD190333
40	275.4	24.8	642	6	BD237072
41	275.4	24.8	642	6	AR225472
42	275.4	24.8	642	6	AX321588
43	274	24.7	905	6	AX337050
44	274	24.7	905	6	BD190295
45	274	24.7	905	9	HS18D

ALIGNMENTS

RESULT 1	AX014337	Sequence 46 from Patent WO9954353.	851 bp	DNA	linear	PAT 07-SEP-2000
AX014337	LOCUS					
DEFINITION	Sequence 46 from Patent WO9954353.					
ACCESSION	AX014337					
VERSION	AX014337.1	GI:10040691				
KEYWORDS	Homo sapiens (human)					
SOURCE	Homo sapiens					
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
REFERENCE	Schmitt,A., Specht,T., Dahl,E., Hinzmann,B., Rosenthal,A. and					
AUTHORS	Pilarsky,C.					
TITLE	Human nucleic acid sequences of normal uterus tissue					

JOURNAL Patent: WO 9954353-A 46 28-OCT-1999;  
 SCHMITT ARMIN (DE); SPECHT THOMAS (DE); DAHL EDGAR (DE); HINZMANN  
 BERND (DE); ROSENTHAL ANDRE (DE); METAGEN GES FUER GENOMFORSCHUN  
 (DE); PILARSKY CHRISTIAN (DE)

FEATURES  
 source  
 Location/Qualifiers  
 1. 851  
 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"

ORIGIN  
 Query Match 72.3%; Score 802; DB 6; Length 851;  
 Best Local Similarity 98.6%; Pred. No. 1.3e-180;  
 Matches 830; Conservative 0; Mismatches 10; Indels 2; Gaps 2;

QY 1 CTTGCACAGAGACACTGGAGGTTTGTAGTCCCAACCCGACAGAGAGGACTGCAGC 60  
 DB 2 CTTGCACAGAGACACTGGAGGTTTGTAGTCCCAACCCGACAGAGAGGACTGCAGC 61  
 QY 61 CTGAGGAAAGAGCAAGGATTTTCAGGAGAGAGGCTCGGACAAAGTGAGCAGGAAATAGAAA 120  
 DB 62 CTGAGGAAAGAGCAAGGATTTTCAGGAGAGAGGCTCGGACAAAGTGAGCAGGAAATAGAAA 121  
 QY 121 CTTAAGAGAAATACACACTTCTGAGAACTGAAACGACAGGGAAGAGGTTCTACTGA 180  
 DB 122 CTTAAGAGAAATACACACTTCTGAGAACTGAAACGACAGGGAAGAGGTTCTACTGA 180  
 QY 181 GCACCGTCCAGCATCCGACACACACAGCGGCTTGTGCTCCAGCGAAACACACTT 240  
 DB 182 CTTAAGAGAAATACACACTTCTGAGAACTGAAACGACAGGGAAGAGGTTCTACTGA 180  
 QY 181 GCACCGTCCAGCATCCGACACACACAGCGGCTTGTGCTCCAGCGAAACACACTT 240  
 DB 182 CTTAAGAGAAATACACACTTCTGAGAACTGAAACGACAGGGAAGAGGTTCTACTGA 180  
 QY 241 CTCACACCTTCTACTCAACACTTCTTCCCAAGAGCAGAGAGATGCAAGGAGGAAATG 300  
 DB 241 CTCACACCTTCTACTCAACACTTCTTCCCAAGAGCAGAGAGATGCAAGGAGGAAATG 300  
 QY 301 AGGTGGCTGTGTGGGGGCAACCCCGACGACCACTTCTTCAAGGTCCACCGTGTATCA 360  
 DB 301 AGGTGGCTGTGTGGGGGCAACCCCGACGACCACTTCTTCAAGGTCCACCGTGTATCA 360

Db 841 GG 842

RESULT 2  
 BD222216  
 LOCUS  
 DEFINITION Human nucleic acid sequence originating in normal uterine tissue.  
 ACCESSION BD222216  
 VERSION BD222216.1 GI:33031986  
 KEYWORDS JP 2002512017-A/45.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCE  
 1 (bases 1 to 851)  
 Specht, T., Hinzmann, B., Schmitt, A., Pilarczyk, C., Dahl, E. and  
 Rosenthal, A.  
 Human nucleic acid sequence originating in normal uterine tissue  
 Patent: JP 2002512017-A 45 23-APR-2002;  
 METAGEN GESELLSCHAFT FUER GENOM FORSCHUNG MBH  
 OS Homo sapiens (human)  
 PN JP 2002512017-A/45  
 PD 23-APR-2002  
 PF 15-APR-1999 JP 2000544691  
 PR 17-APR-1998 DE 198 17 945.4  
 PT THOMAS SPECHT, BERND HINZMANN, ARMIN SCHMITT, CHRISTIAN PILARSKY,  
 PI EDGAR DAHL,  
 PI ANDRE ROSENTHAL  
 PC C12N15/09, A61K38/00, A61K48/00, C07K14/47, C07K16/18,  
 PC C12N1/19,  
 PC C12N1/21, C12N5/00, C12P21/08, C12Q1/68, G01N33/68, C12N15/00, PC  
 A61K37/02,  
 PC C12N5/00  
 CC Human nucleic acid sequence originating in normal uterine CC  
 tissue

FEATURES  
 source  
 1. 851  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"

ORIGIN  
 Query Match 72.3%; Score 802; DB 6; Length 851;  
 Best Local Similarity 98.6%; Pred. No. 1.3e-180;  
 Matches 830; Conservative 0; Mismatches 10; Indels 2; Gaps 2;

QY 1 CTTGCACAGAGACACTGGAGGTTTGTAGTCCCAACCCGACAGAGAGGACTGCAGC 60  
 DB 2 CTTGCACAGAGACACTGGAGGTTTGTAGTCCCAACCCGACAGAGAGGACTGCAGC 61  
 QY 61 CTGAGGAAAGAGCAAGGATTTTCAGGAGAGAGGCTCGGACAAAGTGAGCAGGAAATAGAAA 120  
 DB 62 CTGAGGAAAGAGCAAGGATTTTCAGGAGAGAGGCTCGGACAAAGTGAGCAGGAAATAGAAA 121  
 QY 121 CTTAAGAGAAATACACACTTCTGAGAACTGAAACGACAGGGAAGAGGTTCTACTGA 180  
 DB 122 CTTAAGAGAAATACACACTTCTGAGAACTGAAACGACAGGGAAGAGGTTCTACTGA 180  
 QY 181 GCACCGTCCAGCATCCGACACACACAGCGGCTTGTGCTCCAGCGAAACACACTT 240  
 DB 181 GCACCGTCCAGCATCCGACACACACAGCGGCTTGTGCTCCAGCGAAACACACTT 240  
 QY 241 CTCACACCTTCTACTCAACACTTCTTCCCAAGAGCAGAGAGATGCAAGGAGGAAATG 300  
 DB 241 CTCACACCTTCTACTCAACACTTCTTCCCAAGAGCAGAGAGATGCAAGGAGGAAATG 300  
 QY 301 AGGTGGCTGTGTGGGGGCAACCCCGACGACCACTTCTTCAAGGTCCACCGTGTATCA 360  
 DB 301 AGGTGGCTGTGTGGGGGCAACCCCGACGACCACTTCTTCAAGGTCCACCGTGTATCA 360

```

QY 361 TCCACAGGAGACCTCGGTGCCGACCATCTGCTGGTCCCTCTTCAACACCCCTCTTCT 420
Db 361 TCCACAGGAGACCTCGGTGCCGACCATCTGCTGGTCCCTCTTCAACACCCCTCTTCT 420
QY 421 TGAACCTGGTGTCTGCTGGGCTTTCATAGCATTCGCTACTCCGTGAAAGTCTAGGACAGGA 480
Db 421 TGAACCTGGTGTCTGCTGGGCTTTCATAGCATTCGCTACTCCGTGAAAGTCTAGGACAGGA 480
QY 481 AGATGGTTGGCGACGTGACCGGGGCCAGGCTATGCTCCACCGCAAGTCCCTGAACA 540
Db 481 AGATGGTTGGCGACGTGACCGGGGCCAGGCTATGCTCCACCGCAAGTCCCTGAACA 540
QY 541 TCTGGGCCCTGATTCTGGGCGATCCTCATGACCAATTGGATTTCATCTGTTACTGGTATTCG 600
Db 541 TCTGGGCCCTGATTCTGGGCGATCCTCATGACCAATTGGATTTCATCTGTTACTGGTATTCG 600
QY 601 GCTCTGTGACAGTCTACCATATTTATGTTACAGATAATACAGGAAACCGGGTACTAGT 660
Db 601 GCTCTGTGACAGTCTACCATATTTATGTTACAGATAATACAGGAAACCGGGTACTAGT 660
QY 661 AGCGGCCCATAGCTGCAACCTTTGCACTCAGCTGCAATGCTGGCCCTGCACGCTGGG 720
Db 661 AGCGGCCCATAGCTGCAACCTTTGCACTCAGCTGCAATGCTGGCCCTGCACGCTGGG 720
QY 721 GCTGTGCCCCCTGCCCCCTGGTCTCCCTAGATACAGAGTATTATACCCACACACT 780
Db 721 GCTGTGCCCCCTGCCCCCTGGTCTCCCTAGATACAGAGTATTATACCCACACACT 780
QY 781 GNTACACTGACATTCATATAAGTG-ACGTGCTTGTGAAAAAACAATAAACCAGA 839
Db 781 GNTACACTGACATTCATATAAGTG-ACGTGCTTGTGAAAAAACAATAAACCAGA 839
QY 840 GG 841
Db 841 GG 842

RESULT 3
AX281854 1095 bp DNA linear PAT 02-NOV-2001
LOCUS Sequence 263 from Patent WO0177389.
DEFINITION AX281854
ACCESSION AX281854
VERSION AX281854.1 GI:16609105
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Shiffman,D., Somogyi,R., Lawn,R., Seilhamer,J.J., Porter,G.J.,
Mikita,T. and Tai,J.
TITLE Genes expressed in foam cell differentiation
JOURNAL Patent: WO 0177389-A 263 18-OCT-2001;
Incyte Genomics, Inc. (US)
FEATURES
source
1. .1095
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/notes="Incyte ID No: 988231.7"

ORIGIN
Query Match 63.1%; Score 699.8; DB 6; Length 1095;
Best Local Similarity 99.6%; Pred. No. 3.1e-156;
Matches 712; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 106 AGCAGGAATAGAACTTAAGAGAAATACACATTTCTGAGAAACTGAAACGACAGGGGAA 165
Db 381 AGCAGGAATAGAACTTAAGAGAAATACACATTTCTGAGAAACTGAAACGACAGGGGAA 440
QY 166 AGGAGTCTCAGTGAGCAGCTCCAGCATCCGACACACAGCGGCCCTTGGCTCCAG 225
Db 441 AGGAGTCTCAGTGAGCAGCTCCAGCATCCGACACACAGCGGCCCTTGGCTCCAG 500

```

```

QY 226 CAGAAAACCACTCTCTCAAAACCTTCACTCAACACTTCTCTCCCAAGCCAGAGATGC 285
Db 501 CAGAAAACCACTCTCTCAAAACCTTCACTCAACACTTCTCTCCCAAGCCAGAGATGC 550
QY 286 ACAAGGAGGAACATGAGGTGGCTGTGTGGGGGACACCCCCAGCAGCACCATCTTCCAAGGT 345
Db 561 ACAAGGAGGAACATGAGGTGGCTGTGTGGGGGACACCCCCAGCAGCACCATCTTCCAAGGT 620
QY 346 CACCCGTGATCAATCCACAGGAGACCTCCGTGCCCGACCATCTGCTGTGGTCCCTGT 405
Db 621 CACCCGTGATCAATCCACAGGAGACCTCCGTGCCCGACCATCTGCTGTGGTCCCTGT 680
QY 406 TCAACACCCCTCTTCTTGAACCTGGTGTCTGTGGGCTTTCATAGCATTCGCTACTCCGTGA 465
Db 681 TCAACACCCCTCTTCTTGAACCTGGTGTCTGTGGGCTTTCATAGCATTCGCTACTCCGTGA 740
QY 466 AGTCTAGGACAGAGATGGTTGGCGACGTGACCGGGGCCAGGCTATGCTTCCACCG 525
Db 741 AGTCTAGGACAGAGATGGTTGGCGACGTGACCGGGGCCAGGCTATGCTTCCACCG 800
QY 526 CCAAGTGCCTGAACATCTGGGCCCTGATTCTGGGCATCTCATGACCATTTGATTTCATCC 585
Db 801 CCAAGTGCCTGAACATCTGGGCCCTGATTCTGGGCATCTCATGACCATTTGATTTCATCC 860
QY 586 TGTACTGGTATTGGCTCTGTGACAGTCTACCATATTATTTACAGATAATACAGGAA 645
Db 861 TGTACTGGTATTGGCTCTGTGACAGTCTACCATATTATTTACAGATAATACAGGAA 920
QY 646 AACGGGTTACTAGTAGCCGCCCATAGCTCCCACTTGCACCTCCACTGTGCAATGCTG 705
Db 921 AACGGGTTACTAGTAGCCGCCCATAGCTCCCACTTGCACCTCCACTGTGCAATGCTG 980
QY 706 GCCTGACAGCTGGGGCTGTTGGCCCTGCTCCCTTGGTCTTGGCCCTAGATACAGAGTT 765
Db 981 GCCTGACAGCTGGGGCTGTTGGCCCTGCTCCCTTGGTCTTGGCCCTAGATACAGAGTT 1040
QY 766 TATACCCACACACTGTCTACACTGACATTCATTAAGTG-ACGTGCTTGTGAAA 819
Db 1041 TATACCCACACACTGTCTACACTGACATTCATTAAGTG-ACGTGCTTGTGAAA 1095

RESULT 4
AX397528 842 bp DNA linear PAT 18-MAY-2002
LOCUS Sequence 1743 from Patent WO0212328.
DEFINITION AX397528
ACCESSION AX397528
VERSION AX397528.1 GI:21068275
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS King,G.E., Meagher,M.J., Xu,J. and Secrist,H.
TITLE Compositions and methods for the therapy and diagnosis of colon
JOURNAL Patent: WO 0212328-A 1743 14-FEB-2002;
CORIXA CORPORATION (US)
FEATURES
source
1. .842
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 62.8%; Score 696.2; DB 6; Length 842;
Best Local Similarity 99.4%; Pred. No. 2.3e-155;
Matches 709; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 106 AGCAGGAATAGAACTTAAGAGAAATACACATTTCTGAGAAACTGAAACGACAGGGGAA 165
Db 130 AGCAGGAATAGAACTTAAGAGAAATACACATTTCTGAGAAACTGAAACGACAGGGGAA 189

```

166 AGGAGGTTCTCAGTGGACCGTCCAGCATCCGAGATCCGAGACCCAGCGCCCTTCGCTCCACG 225  
 190 AGGAGGTTCTCAGTGGACCGTCCAGCATCCGAGATCCGAGACCCAGCGCCCTTCGCTCCACG 249  
 226 CAGAAACCAACATCTTCAAAACCTTCACTCAACATCTTCCCTCCCAAGCCAGAGATGC 295  
 250 CAGAAACCAACATCTTCAAAACCTTCACTCAACATCTTCCCTCCCAAGCCAGAGATGC 309  
 286 ACAAGGAGGAACATGAGTGGCTGTGCTGGGGGACCCCGCCAGCAGCATCTTCCAAAGGT 345  
 310 ACAAGGAGGAACATGAGTGGCTGTGCTGGGGGACCCCGCCAGCAGCATCTTCCAAAGGT 369  
 346 CCACCGTGATCAACATCCACAGCAGACCTCCGTCGCGCCGACCATGTCGTCGCTCCCTGT 405  
 370 CCACCGTGATCAACATCCACAGCAGACCTCCGTCGCGCCGACCATGTCGTCGCTCCCTGT 429  
 406 TCAACACCCCTCTTCTTGAACCTGGTCTGTCTGGGCTTCATAGCATTCGCGCTACTCCGTGA 465  
 430 TCAACACCCCTCTTCTTGAACCTGGTCTGTCTGGGCTTCATAGCATTCGCGCTACTCCGTGA 489  
 466 AGTCTAGGACAGGAAGATGGTTGGCAGCTGACCGGGGCCAGGCGCTATGCTCCACCG 525  
 490 AGTCTAGGACAGGAAGATGGTTGGCAGCTGACCGGGGCCAGGCGCTATGCTCCACCG 549  
 526 CCAAGTCCCTGAACATCTGGGCGCTGATTCGCGCATCTCATGACCATTTGGATTTCATCC 585  
 550 CCAAGTCCCTGAACATCTGGGCGCTGATTCGCGCATCTCATGACCATTTGGATTTCATCC 609  
 586 TGTTACTGGTATTCGGCTCTGTGACAGCTACCATATTTATTTACAGATAAATACAGGAAA 645  
 610 TGTCACTGGTATTCGGCTCTGTGACAGCTACCATATTTATTTACAGATAAATACAGGAAA 669  
 646 AACGGGTTACTAGTAGCGGCCATGAGCTGCAACTTTGGCTCCACTCCAGTGCATGCTG 705  
 670 AACGGGTTACTAGTAGCGGCCATGAGCTGCAACTTTGGCTCCACTCCAGTGCATGCTG 729  
 706 GCCCTGCACGCTGGGGCTGTGCTCCCTGCGCCCTTGGTCTGCTCCCTAGATACAGAGTT 765  
 730 GCCCTGCACGCTGGGGCTGTGCTCCCTGCGCCCTTGGTCTGCTCCCTAGATACAGAGTT 789  
 766 TATACCCACACACCTGTCTACAGTGTCTATCAATTAAGTGCAGCTGCTGTGA 817  
 790 TATACCCACACACCTGTCTACAGTGTCTATCAATTAAGTGCAGCTGCTGTGA 842

RESULT 5  
 AX302539  
 LOCUS AX302539  
 DEFINITION Sequence 57 from Patent WO0175177.  
 ACCESSION AX302539  
 VERSION AX302539.1 GI:17383079  
 KEYWORDS Homo sapiens (human)  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1  
 AUTHORS Morin,P.J., Sherman-Baust,C.A., Pizer,E.S. and Hough,C.D.  
 TITLE Tumor markers in ovarian cancer  
 JOURNAL Patent: WO 0175177-A 57 11-OCT-2001;  
 THE SECRETARY OF THE DEPARTMENT OF HEALTH AND HUMAN SERVICES (US)  
 FEATURES  
 Location/Qualifiers  
 1..853  
 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"  
 ORIGIN  
 Query Match 62.8%; Score 696.2; DB 6; Length 853;  
 Best Local Similarity 99.4%; Pred. No. 2.3e-155;  
 Matches 709; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

106 AGCAGGAAATAGAAACCTTAAGAGAAATACACACTTCTGAGAAATGAAACGACAGGGGAA 165  
 141 AGCAGGAAATAGAAACCTTAAGAGAAATACACACTTCTGAGAAATGAAACGACAGGGGAA 200  
 166 AGGAGGTTCTCAGTGGACCGTCCAGCATCCGAGACCCAGCGCCCTTCGCTCCACG 225  
 201 AGGAGGTTCTCAGTGGACCGTCCAGCATCCGAGACCCAGCGCCCTTCGCTCCACG 260  
 226 CAGAAACCAACATCTTCAAAACCTTCACTCAACATCTTCCCTCCCAAGCCAGAGATGC 285  
 261 CAGAAACCAACATCTTCAAAACCTTCACTCAACATCTTCCCTCCCAAGCCAGAGATGC 320  
 286 ACAAGGAGGAACATGAGTGGCTGTGCTGGGGGACCCCGCCAGCAGCATCTTCCAAAGGT 345  
 321 ACAAGGAGGAACATGAGTGGCTGTGCTGGGGGACCCCGCCAGCAGCATCTTCCAAAGGT 380  
 346 CCACCGTGATCAACATCCACAGCAGACCTCCGTCGCGCCGACCATGTCGTCGCTCCCTGT 405  
 381 CCACCGTGATCAACATCCACAGCAGACCTCCGTCGCGCCGACCATGTCGTCGCTCCCTGT 440  
 406 TCAACACCCCTCTTCTTGAACCTGGTCTGTCTGGGCTTCATAGCATTCGCGCTACTCCGTGA 465  
 441 TCAACACCCCTCTTCTTGAACCTGGTCTGTCTGGGCTTCATAGCATTCGCGCTACTCCGTGA 500  
 466 AGTCTAGGACAGGAAGATGGTTGGCAGCTGACCGGGGCCAGGCGCTATGCTCCACCG 525  
 501 AGTCTAGGACAGGAAGATGGTTGGCAGCTGACCGGGGCCAGGCGCTATGCTCCACCG 560  
 526 CCAAGTCCCTGAACATCTGGGCGCTGATTCGCGCATCTCATGACCATTTGGATTTCATCC 585  
 561 CCAAGTCCCTGAACATCTGGGCGCTGATTCGCGCATCTCATGACCATTTGGATTTCATCC 620  
 586 TGTTACTGGTATTCGGCTCTGTGACAGCTACCATATTTATTTACAGATAAATACAGGAAA 645  
 621 TGTCACTGGTATTCGGCTCTGTGACAGCTACCATATTTATTTACAGATAAATACAGGAAA 680  
 646 AACGGGTTACTAGTAGCGGCCATGAGCTGCAACTTTGGCTCCACTCCAGTGCATGCTG 705  
 681 AACGGGTTACTAGTAGCGGCCATGAGCTGCAACTTTGGCTCCACTCCAGTGCATGCTG 740  
 706 GCCCTGCACGCTGGGGCTGTGCTCCCTGCGCCCTTGGTCTGCTCCCTAGATACAGAGTT 765  
 741 GCCCTGCACGCTGGGGCTGTGCTCCCTGCGCCCTTGGTCTGCTCCCTAGATACAGAGTT 800  
 766 TATACCCACACACCTGTCTACAGTGTCTATCAATTAAGTGCAGCTGCTGTGA 817  
 801 TATACCCACACACCTGTCTACAGTGTCTATCAATTAAGTGCAGCTGCTGTGA 853

RESULT 6  
 AX409461  
 LOCUS AX409461  
 DEFINITION Sequence 2108 from Patent WO0229103.  
 ACCESSION AX409461  
 VERSION AX409461.1 GI:21442166  
 KEYWORDS Homo sapiens (human)  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1  
 AUTHORS Alvares,C., Horne,D., Peres-da-Silva,S. and Vockley,J.G.  
 TITLE Gene expression profiles in liver cancer  
 JOURNAL Patent: WO 0229103-A 2108 11-APR-2002;  
 GENE LOGIC INC (US)  
 FEATURES  
 Location/Qualifiers  
 1..853  
 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"  
 /note="EMBL/GenBank Accession No. J04164"  
 ORIGIN

```

Query Match      62.8%; Score 696.2; DB 6; Length 853;
Best Local Similarity 99.4%; Pred. No. 2.3e-155;
Matches 709; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 106 AGCAGGAAATAGAAATTAAGAGAAATACACACTTCTGAGAAATCGAAACGACAGGGGAA 165
DB 141 AGCAGGAAATAGAAATTAAGAGAAATACACACTTCTGAGAAATCGAAACGACAGGGGAA 200

QY 166 AGGAGTCTCACTGAGACCGTCCAGCATCCGACACACACAGCGGCCCTTGCCTCCAG 225
DB 201 AGGAGTCTCACTGAGACCGTCCAGCATCCGACACACACAGCGGCCCTTGCCTCCAG 260

QY 226 CAGAAAACCACTTCTCAAACTTCACTCAACACTTCTTCCCCAAAGCCAGAGAGATGC 285
DB 261 CAGAAAACCACTTCTCAAACTTCACTCAACACTTCTTCCCCAAAGCCAGAGAGATGC 320

QY 286 ACNAGAGGAAACATGAGTGGCTGTCTGGGGGCAACCCGACGACCACTTCCAGGT 345
DB 321 ACNAGAGGAAACATGAGTGGCTGTCTGGGGGCAACCCGACGACCACTTCCAGGT 380

QY 346 CCACCGTGATCAACATCCACAGGAGACCTCCGTGCGGACCATCTCGTCTGGTCCCTGT 405
DB 381 CCACCGTGATCAACATCCACAGGAGACCTCCGTGCGGACCATCTCGTCTGGTCCCTGT 440

QY 406 TCAACACCCCTCTTCTTGAACCTGTGTCTGGGGTTCATAGCATTTCGCCTACTCCGTGA 465
DB 441 TCAACACCCCTCTTCTTGAACCTGTGTCTGGGGTTCATAGCATTTCGCCTACTCCGTGA 500

QY 466 AGTCTAGGACAGGAAGATGGTGGGAGCGTGACCCGGGGCCAGGCGCTATGCTCCACCG 525
DB 501 AGTCTAGGACAGGAAGATGGTGGGAGCGTGACCCGGGGCCAGGCGCTATGCTCCACCG 560

QY 526 CCAAGTCCCTGAAACATCTGGGCCCTGATTCTGGGCATCTCTGACCATTCGATTCATCC 585
DB 561 CCAAGTCCCTGAAACATCTGGGCCCTGATTCTGGGCATCTCTGACCATTCGATTCATCC 620

QY 586 TGTACTGGTATTTCGGTCTCTGTGACAGCTTACCATATTATGTTACAGATAATACAGAAA 645
DB 621 TGTACTGGTATTTCGGTCTCTGTGACAGCTTACCATATTATGTTACAGATAATACAGAAA 680

QY 646 AACGGGGTACTAGTACGCGCCCATAGCTGCAACCTTTGCACTCCACTGTGCAATCTGT 705
DB 681 AACGGGGTACTAGTACGCGCCCATAGCTGCAACCTTTGCACTCCACTGTGCAATCTGT 740

QY 706 GCCTTGACGCTGGGGTGTGGTCCCTGCGCCCTGGTCTGCTCCCTAGATACAGAGTT 765
DB 741 GCCTTGACGCTGGGGTGTGGTCCCTGCGCCCTGGTCTGCTCCCTAGATACAGAGTT 800

QY 766 TATACCCACACACCTGTCTACACTGACATTCAATAAAGTG-ACGTGCTGTGTA 817
DB 801 TATACCCACACACCTGTCTACAGTGTCAATCAATAAAGTGACAGTGTGTGTA 853

RESULT 7
LOCUS HUM927A
DEFINITION Human interferon-inducible protein 9-27 mRNA, complete cds.
ACCESSION J04164
VERSION J04164.1 GI:177801
KEYWORDS interferon-inducible protein.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 853)
AUTHORS Reid,L.E., Brasnett,A.H., Gilbert,C.S., Porter,A.C., Gewert,D.R.,
Starck,G.N. and Kerr,I.M.
TITLE A single DNA response element can confer inducibility by both
alpha- and gamma-interferons
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 86 (3), 840-844 (1989)
MEDLINE 89128873
PUBMED 2492664
COMMENT Original source text: Human lymphoid cell, cdNA to mRNA, and DNA,

```

```

clones php[1,7,2,0,3,4].
Computer-readable sequence for [1] kindly provided by I.M.Kerr,
07-FEB-1989.
Bases 1-178 are of genomic origin. An intron is located between
positions 501 and 502 in the genomic DNA.
FEATURES
Location/Qualifiers
source 1..853
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
prim_transcript 177..852
/note="9-27 mRNA"
prim_transcript 205..852
/note="9-27 mRNA"
CDS 317..694
/note="interferon-inducible protein 9-27"
/codon_start=1
/protein_id="AAA35494.1"
/db_xref="GI:177802"
/translation="MKHEHEVAVLGAPESTILPRSTVINIHSETSPVDHVMWLSFNT
LFLNWCCLGFIAFAYSVKSRDRKXMDVDTGAQAVASTAKLNIALILGILMTIGFIL
SLVFGSVTVVHIMLQIQEKG"
ORIGIN 156 bp upstream of AflII site.

Query Match      62.8%; Score 696.2; DB 9; Length 853;
Best Local Similarity 99.4%; Pred. No. 2.3e-155;
Matches 709; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 106 AGCAGGAAATAGAAATTAAGAGAAATACACACTTCTGAGAAATCGAAACGACAGGGGAA 165
DB 141 AGCAGGAAATAGAAATTAAGAGAAATACACACTTCTGAGAAATCGAAACGACAGGGGAA 200

QY 166 AGGAGTCTCACTGAGACCGTCCAGCATCCGACACACACAGCGGCCCTTGCCTCCAG 225
DB 201 AGGAGTCTCACTGAGACCGTCCAGCATCCGACACACACAGCGGCCCTTGCCTCCAG 260

QY 226 CAGAAAACCACTTCTCAAACTTCACTCAACACTTCTTCCCCAAAGCCAGAGAGATGC 285
DB 261 CAGAAAACCACTTCTCAAACTTCACTCAACACTTCTTCCCCAAAGCCAGAGAGATGC 320

QY 286 ACAAGGAGGAAACATGAGTGGCTGTGTCTGGGGCACCCGCCAGCACCATCTTCCAAGGT 345
DB 321 ACAAGGAGGAAACATGAGTGGCTGTGTCTGGGGCACCCGCCAGCACCATCTTCCAAGGT 380

QY 346 CCACCGTGATCAACATCCACAGGAGACCTCCGTGCGGACCATCTCGTCTGGTCCCTGT 405
DB 381 CCACCGTGATCAACATCCACAGGAGACCTCCGTGCGGACCATCTCGTCTGGTCCCTGT 440

QY 406 TCAACACCCCTCTTCTTGAACCTGTGTCTGGGGTTCATAGCATTTCGCCTACTCCGTGA 465
DB 441 TCAACACCCCTCTTCTTGAACCTGTGTCTGGGGTTCATAGCATTTCGCCTACTCCGTGA 500

QY 466 AGTCTAGGACAGGAAGATGGTGGGAGCGTGACCCGGGGCCAGGCGCTATGCTCCACCG 525
DB 501 AGTCTAGGACAGGAAGATGGTGGGAGCGTGACCCGGGGCCAGGCGCTATGCTCCACCG 560

QY 526 CCAAGTCCCTGAAACATCTGGGCCCTGATTCTGGGCATCTCTGACCATTCGATTCATCC 585
DB 561 CCAAGTCCCTGAAACATCTGGGCCCTGATTCTGGGCATCTCTGACCATTCGATTCATCC 620

QY 586 TGTACTGGTATTTCGGTCTCTGTGACAGCTTACCATATTATGTTACAGATAATACAGAAA 645
DB 621 TGTACTGGTATTTCGGTCTCTGTGACAGCTTACCATATTATGTTACAGATAATACAGAAA 680

QY 646 AACGGGGTACTAGTACGCGCCCATAGCTGCAACCTTTGCACTCCACTGTGCAATCTGT 705
DB 681 AACGGGGTACTAGTACGCGCCCATAGCTGCAACCTTTGCACTCCACTGTGCAATCTGT 740

QY 706 GCCTTGACGCTGGGGTGTGGTCCCTGCGCCCTGGTCTGCTCCCTAGATACAGAGTT 765
DB 741 GCCTTGACGCTGGGGTGTGGTCCCTGCGCCCTGGTCTGCTCCCTAGATACAGAGTT 800

QY 766 TATACCCACACACCTGTCTACACTGACATTCAATAAAGTG-ACGTGCTGTGTA 817

```



```
Db 801 TATACCCACACCTGCTACAGTGTCAATTAATAAGTGACGCTGTGTGA 853
|||||
BC000897 683 bp mRNA linear PRI 12-NOV-2003
LOCUS BC000897
DEFINITION Homo sapiens interferon induced transmembrane protein 1 (9-27),
ACCESSION BC000897
VERSION BC000897.1 GI:12654158
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 683)
AUTHORS Klausner, R.D., Collins, F.S., Wagner, L.H., Derge, J.G.,
Mammalia, Euthera; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 683)
Strausberg, R.L., Feigold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L.H., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.P., Casavant, T.L.,
Scheetz, T.E., Brownstein, M.J., Udgin, T.B., Tohiyuki, S.,
Carninci, P., Frange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalon, D.K., Munz, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shavchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smallos, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16999-16903 (2002)
22388257
12477932
2 (bases 1 to 683)
Strausberg, R.
Direct Submission
Submitted (15-NOV-2000) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgabbs@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: http://www-shgc.stanford.edu
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAK Plate: 4 Row: b Column: 24
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 4504580.
Location/Qualifiers
1. .683
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:5195 IMAGE:3464598"
/tissue_type="Cervix, carcinoma"
/clone_lib="NIH_MGC_12"
FEATURES
source
```

```
/lab_host="DH10B"
/notes="Vector: pCMV-SPORT6"
1. .683
/genes="IFITM1"
/notes="Synonyms: 9-27, IFI17, LEU13, CD225"
/db_xref="LocusID:8519"
/db_xref="MIM:604456"
131..508
/codon_start=1
/product="interferon induced transmembrane protein 1
(9-27)"
/protein_id="AAH00897.1"
/db_xref="GI:12654159"
/db_xref="LocusID:8519"
/translation="MKKEHEVAVIGAPPSTILPRSTVINIHSETSVDPHVVSLPNT
LLNWCCLGFIAPYSVKSDRMVGDVGTGAQVASTAKCLINAILILGILMTIGFIL
LLVFGSVTVYHIMLQIIQKRGY"
188..433
/notes="CD225; Region: Interferon-induced transmembrane
protein. This family includes the human leukocyte antigen
CD225, which is an interferon inducible transmembrane
protein, and is associated with interferon induced cell
growth suppression"
/db_xref="CDD:pfam04505"
misc_feature
ORIGIN
Query Match 59.9%; Score 664.6; DB 9; Length 683;
Best Local Similarity 99.3%; Pred. No. 8.1e-148;
Matches 678; Conservative 0; Mismatches 4; Indels 1; Gaps 1;
QY 152 AAAGCAGCGGGGAAAGAGGTCTCAGTGAGCAGCGTCCAGCATCCGAGCAGCAGCAGCGG 211
DB 1 AAAGCAGCGGGGAAAGAGGTCTCAGTGAGCAGCGTCCAGCATCCGAGCAGCAGCAGCGG 60
QY 212 CCCTTCGCTCCAGCGAGAAACACACTTCTCAAACTTCACTCAACTTCTTCTTCCCA 271
DB 61 CCCTTCGCTCCAGCGAGAAACACACTTCTCAAACTTCACTCAACTTCTTCTTCCCA 120
QY 272 AAGCCAGAGATGACAGAGGAGGAGCATGAGTGGTGTCTGGGGGACCCCGCAGCAC 331
DB 121 AAGCCAGAGATGACAGAGGAGGAGCATGAGTGGTGTCTGGGGGACCCCGCAGCAC 180
QY 332 CATCTTCCAAAGTCCACCGTATCAACATCCAGCGAGAGCTCCGTCGCGCAGCATGT 391
DB 181 CATCTTCCAAAGTCCACCGTATCAACATCCAGCGAGAGCTCCGTCGCGCAGCATGT 240
QY 392 CGTCTGGTCCCTGTTCAACACCCCTCTTCTTGAAGTGGTGTCTGTGGGCTTCATAGCAT 451
DB 241 CGTCTGGTCCCTGTTCAACACCCCTCTTCTTGAAGTGGTGTCTGTGGGCTTCATAGCAT 300
QY 452 CGCCTACTCCGTGAAGTCTAGGAGCAGGAGAGATGGTTGGCGAGCTGACCGGGGCGCAGGC 511
DB 301 CGCCTACTCCGTGAAGTCTAGGAGCAGGAGAGATGGTTGGCGAGCTGACCGGGGCGCAGGC 360
QY 512 CTATGCTCCACCGCGAAGTCCGTGACATCTGGCCCTGATTCTGGGCATCTCATGAC 571
DB 361 CTATGCTCCACCGCGAAGTCCGTGACATCTGGCCCTGATTCTGGGCATCTCATGAC 420
QY 572 CATTTGATTTCATCTCTGTTACTGTTATTTGGGCTCTGTGACAGTCTACCATATATTATGTACA 631
DB 421 CATTTGATTTCATCTCTGTTACTGTTATTTGGGCTCTGTGACAGTCTACCATATATTATGTACA 480
QY 632 GATATACAGGAGAAACGGGTTACTAGTAGCCGCCCATAGCTGACACCTTTGACATCC 691
DB 481 GATATACAGGAGAAACGGGTTACTAGTAGCCGCCCATAGCTGACACCTTTGACATCC 540
QY 692 ACTGTGCAATGTGGCCCTGTCAGCGTGGGGCTGTGTGCCCTTCCCTTGGTCTGGCC 751
DB 541 ACTGTGCAATGTGGCCCTGTCAGCGTGGGGCTGTGTGCCCTTCCCTTGGTCTGGCC 600
QY 752 TAGATACAGAGTTTATACCCACACACTCTCTACACTGACATTCATTAAGTACAGTG 810
DB 601 TAGATACAGAGTTTATACCCACACACTCTCTACAGTGTCTCATTAATAAGTACAGTG 660
```

```
QY      811  CTTGTGAAAAAATAAAATATAA 833
Db      661  CTTGTGAAAAAATAAAATATAA 683

RESULT 9
HSR927
LOCUS   HSR927          647 bp  mRNA  linear  PRI 17-FEB-1997
DEFINITION H.sapiens mRNA for interferon-induced 17kDa membrane protein.
ACCESSION X84958
VERSION   X84958.1 GI:1177475
KEYWORDS  9-27 gene; interferon-inducible protein; Leu-13 gene.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
AUTHORS  Deblandre, G.A., Marinx, O.P., Evans, S.S., Majja, S., Leo, O.,
           Caput, D., Huez, G.A. and Wathelot, M.G.
TITLE     Expression cloning of an interferon-inducible 17-kDa membrane
           protein implicated in the control of cell growth
JOURNAL   J. Biol. Chem. 270 (40), 23860-23866 (1995)
MEDLINE   96007544
PUBMED    7559564
REFERENCE 2 (bases 1 to 647)
AUTHORS  Deblandre, G.A.
TITLE     Direct Submission
JOURNAL   Submitted (22-FEB-1995) G.A. Deblandre, Universite Libre de
           Bruxelles, Departement de Biologie Moleculaire, 67, rue des
           Chevaux, B-1640 Rhode-St-Genese, BELGIUM
COMMENT   Sequence overlapping with that under the acc# J04164.
FEATURES
         source
         1..647
         /organism="Homo sapiens"
         /mol_type="mRNA"
         /db_xref="taxon:9606"
         /cell_line="Hela"
         /clone_lib="ATCC CCL2"
         1..647
         /gene="9-27/Leu-13"
         74
         /gene="9-27/Leu-13"
         /note="J04164"
         /replace="a"
         111..488
         /gene="9-27/Leu-13"
         /codon_start=1
         /product="interferon-inducible protein"
         /protein_id="CAA59337.1"
         /db_xref="GI:1177476"
         /db_xref="GOA:P13164"
         /db_xref="SWISS-PROT:P13164"
         /translation="MKKEHEVAVLGAPPSTILPRSTVINIHSTSVDPHVWVSUFTN
         LFNWCCLGFIAPFAYSKSRDRKMGVDVTGAQASTAKCLNIALILGILMTIGFIL
         LLYFGSVTVYHIMLIQIKRGY"
         185
         /gene="9-27/Leu-13"
         /note="J04164"
         /replace="c"
         293
         /gene="9-27/Leu-13"
         /note="J04164"
         /replace="g"
         418
         /gene="9-27/Leu-13"
         /note="J04164"
         /replace="c"
         437
         /gene="9-27/Leu-13"
         /note="J04164"
         /replace="g"
         626..631
         polyA_signal
```

```
ORIGIN
Query Match      56.3%; Score 624.4; DB 9; Length 647;
Best Local Similarity 98.9%; Pred. No. 3.2e-138;
Matches 639; Conservative 0; Mismatches 6; Indels 1; Gaps 1;
/ gene="9-27/Leu-13"

QY      173  CTCACTGAGCAGCCCTCCAGCATCCGGACACACACAGCGCGCCCTTCCTCCACCCAGAAA 232
Db      2    CTCACTGAGCAGCCCTCCAGCATCCGGACACACACAGCGCGCCCTTCCTCCACCCAGAAA 61

QY      233  CCACACTTCTCAAAACCTTCACTCAACACTTCTTCCCAAGCCAGAAAGATGCACAAAGGA 292
Db      62  CCACACTTCTCATACCTTCACTCAACACTTCTTCCCAAGCCAGAAAGATGCACAAAGGA 121

QY      293  GGAACATGAGTGCGTGTCTGGGGGACCCCGACACCATCTTCCAGGTTCCACCGT 352
Db      122  GGAACATGAGTGCGTGTCTGGGGGACCCCGACACCATCTTCCAGGTTCCACCGT 181

QY      353  GATCAACATCCACAGCGAGAGCTCCGTCGCCGACCATGTCGTCTGCTCCCTGTTCAACAC 412
Db      182  GATTAACATCCACAGCGAGAGCTCCGTCGCCGACCATGTCGTCTGCTCCCTGTTCAACAC 241

QY      413  CCTCTTCTTGAACGTGCTGTCTGGGCTTCATAGCATTCGCTACTCGTGAAGTCTAG 472
Db      242  CCTCTTCTTGAACGTGCTGTCTGGGCTTCATAGCATTCGCTACTCGTGAAGTCTAG 301

QY      473  GGACAGAAAGATGTTGGCGACGTGACCGGGGCCAGGCGCTATGCTCCACCCCAAGTG 532
Db      302  GGACAGAAAGATGTTGGCGACGTGACCGGGGCCAGGCGCTATGCTCCACCCCAAGTG 361

QY      533  CTGGAACATCTGGGCCCTGATTCGGGCATCTCATAGCATTCGCTACTCGTGAAGTCTAG 592
Db      362  CTGGAACATCTGGGCCCTGATTCGGGCATCTCATAGCATTCGCTACTCGTGAAGTCTAG 421

QY      593  GSTATTGGCTCTGTGACAGTCTACCATATTATTTACAGATAATACAGAAAAACGGGG 652
Db      422  GSTATTGGCTCTGTATACAGTCTACCATATTATTTACAGATAATACAGAAAAACGGGG 481

QY      653  TTACTAGTAGCGGCCATAGCTGCAACCTTTGCACTCCACGTGTGCAATGTCGCGCCCTGC 712
Db      482  TTACTAGTAGCGGCCATAGCTGCAACCTTTGCACTCCACGTGTGCAATGTCGCGCCCTGC 541

QY      713  ACGTGGGGCTGTTGCCCTCCCTGCTGCTCCCTAGATACAGAGTTTATACCC 772
Db      542  ACGTGGGGCTGTTGCCCTCCCTGCTGCTCCCTAGATACAGAGTTTATACCC 601

QY      773  ACACACCTGTCTACACTGACATTCATATAAAGTG-ACGTGCTTGTGA 817
Db      602  ACACACCTGTCTACAGTGTCTATCAATAAAGTGACAGTGTCTTGTGA 647

RESULT 10
BT007173
LOCUS   BT007173          378 bp  mRNA  linear  PRI 13-MAY-2003
DEFINITION Homo sapiens interferon induced transmembrane protein 1 (9-27)
           mRNA, complete cds.
ACCESSION BT007173
VERSION   BT007173.1 GI:30593184
KEYWORDS  FLI_CDNA.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 378)
AUTHORS  Kalnine, N., Chen, X., Rolfs, A., Halleck, A., Hines, L., Eisenstein, S.,
           Koundinya, M., Raphael, J., Moreira, D., Kelley, T., LaBaer, J., Lin, Y.,
           Phelan, M. and Farmer, A.
TITLE     Cloning of human full-length CDSs in BD Creator(TM) System Donor
           vector
JOURNAL   Unpublished
REFERENCE 2 (bases 1 to 378)
AUTHORS  Kalnine, N., Chen, X., Rolfs, A., Halleck, A., Hines, L., Eisenstein, S.,
```

Koundinya.M., Raphael,J., Moreira,D., Kelley,T., Labaer,J., Lin,Y.,  
Phelan,M. and Farmer,A.

# TITLE JOURNAL

Submitted (13-MAY-2003) BD Biosciences Clontech, 1020 East Meadow  
Circle, Palo Alto, CA 94303, USA

## COMMENT

This CDS clone is a part of a collection of human full length  
expression clones generated by BD Biosciences Clontech and the  
Harvard Institute of Proteomics. Each CDS has been cloned in two  
forms: with and without stop-codon (to allow fusion with C-terminal  
tag). The CDS has been directionally cloned using BD In-Fusion(TM)  
cloning system between the SalI and HindIII sites of the pDNR-DUAL  
vector. Additional sequences in the clone: 'ACC' after SalI site  
and before 'ATG' to provide Kozak consensus sequence; 'GG' after  
last codon and before HindIII site to maintain reading frame.  
Clone distribution: http://bioinfo.clontech.com/orfclones.

## FEATURES

source

1..378  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="GH00049X1.0"  
/clone\_lib="BD Creator(TM) CDS Library derived from MGC  
collection"  
/lab\_host="DH5alpha T1 resistant"  
/note="Vector: pDNR-Dual"  
1..378  
/product="interferon induced transmembrane protein 1  
(9-27)"  
/protein\_id="AAP35837.1"  
/translation="MHKEHEVAVLGPSTILPRSTVINIHSETSVDPHVWLSLNT  
LFLNWCCLGFTAFAYSVKSRDRKMGVDVTGAQAYASTAKCLNIWALILGILMTIGFIL  
LLVFGSVTVYHIMLQIIQEKRGY"

## CDS

1..378  
/codon\_start=1  
/product="interferon induced transmembrane protein 1  
(9-27)"  
/protein\_id="AAP35837.1"  
/translation="MHKEHEVAVLGPSTILPRSTVINIHSETSVDPHVWLSLNT  
LFLNWCCLGFTAFAYSVKSRDRKMGVDVTGAQAYASTAKCLNIWALILGILMTIGFIL  
LLVFGSVTVYHIMLQIIQEKRGY"

## ORIGIN

Query Match 34.1%; Score 378; DB 9; Length 378;  
Best Local Similarity 100.0%; Pred. No. 2.1e-79;  
Matches 378; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 282 ATGCACAGGAGGAACATGAGGTGGCTGTCTGGGGGACCCCGCCAGCACCCTTCCA 341  
DB 1 ATGCACAGGAGGAACATGAGGTGGCTGTCTGGGGGACCCCGCCAGCACCCTTCCA 60  
QY 342 AGGTCCACCGTATCAACATCCACAGCAGACCTCGTGGCCGACCATGTCGTGTCTCC 401  
DB 61 AGGTCCACCGTATCAACATCCACAGCAGACCTCGTGGCCGACCATGTCGTGTCTCC 120  
QY 402 CTGTTCAACACCCCTCTTCTTGAACCTGCTGTCTGGCTTTCATAGCATTCGCTACTCC 461  
DB 121 CTGTTCAACACCCCTCTTCTTGAACCTGCTGTCTGGCTTTCATAGCATTCGCTACTCC 180  
QY 462 GTGAAGTCTAGGACAGGAGATGGTTGGCGACGTGACCGGGGCCAGGCTATGCTCTCC 521  
DB 181 GTGAAGTCTAGGACAGGAGATGGTTGGCGACGTGACCGGGGCCAGGCTATGCTCTCC 240  
QY 522 ACCGCCAGTGTCTGAACATCTGGGCCCTGATTCGGGATCTCTATGACCATTTGATTC 581  
DB 241 ACCGCCAGTGTCTGAACATCTGGGCCCTGATTCGGGATCTCTATGACCATTTGATTC 300  
QY 582 ATCTGTTTACTGGTATTCGGCTCTGTGACAGTCTACCATATTATGTTACAGATAATACAG 641  
DB 301 ATCTGTTTACTGGTATTCGGCTCTGTGACAGTCTACCATATTATGTTACAGATAATACAG 360  
QY 642 GAAAACGGGTACTAG 659  
DB 361 GAAAACGGGTACTAG 378

## RESULT 11

BT007946

## LOCUS

DEFINITION Synthetic construct Homo sapiens interferon induced transmembrane

## ACCESSION

BT007946

## VERSION

BT007946.1

## KEYWORDS

FLI CDNA.

## SOURCE

synthetic construct  
artificial construct

## ORGANISM

1 (bases 1 to 378)

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## FEATURES

## source

## ORIGIN

## Query Match

## Best Local Similarity

## Matches

## Conservative

## Mismatches

## Indels

## Gaps

## CDS

## ORIGIN

## Query Match

## Best Local Similarity

## Matches

## Conservative

## Mismatches

## Indels

## Gaps

## CDS

## ORIGIN

## Query Match

## Best Local Similarity

## Matches

## Conservative

## Mismatches

## Indels

## Gaps

## CDS

## ORIGIN

## Query Match

## Best Local Similarity

## Matches

## Conservative

## Mismatches

## Indels

## Gaps

## CDS

## ORIGIN

## Query Match

## Best Local Similarity

## Matches

## Conservative

## Mismatches

## Indels

## Gaps

## CDS

## ORIGIN

## Query Match

## Best Local Similarity

## Matches

## Conservative

## Mismatches

## Indels

## Gaps

## CDS

## ORIGIN

## Query Match

## Best Local Similarity

## Matches

## Conservative

## Mismatches

## Indels

## Gaps

## CDS

## ORIGIN

## Query Match

## Best Local Similarity

## Matches

## Conservative

## Mismatches

## Indels

## Gaps

## CDS

## ORIGIN

## Query Match

## Best Local Similarity

## Matches

## Conservative

## Mismatches

## Indels

## Gaps

## CDS

## ORIGIN

## Query Match

## Best Local Similarity

## Matches

## Conservative

## Mismatches

## Indels

## Gaps

## CDS

## ORIGIN

## Query Match

## Best Local Similarity

## Matches

## Conservative

## Mismatches

## Indels

## Gaps

## CDS

## ORIGIN

## Query Match

## Best Local Similarity

## Matches

## Conservative

## Mismatches

## Indels

## Gaps

## CDS

## ORIGIN

## Query Match

## Best Local Similarity

## Matches

## Conservative

## Mismatches

## Indels

## Gaps

## CDS

## ORIGIN

## Query Match

## Best Local Similarity

## Matches

## Conservative

## Mismatches

## Indels

## Gaps

## CDS

## ORIGIN

## Query Match

## Best Local Similarity

## Matches

## Conservative

## Mismatches

## Indels

## Gaps

## CDS

## ORIGIN

## Query Match

## Best Local Similarity

## Matches

## Conservative

## Mismatches

## Indels

## Gaps

## CDS

## ORIGIN

## Query Match

## Best Local Similarity

## Matches

## Conservative

## Mismatches

## Indels

## Gaps

## CDS

## ORIGIN

## Query Match

## Best Local Similarity

## Matches

## Conservative

## Mismatches

## Indels

## Gaps

## CDS

## ORIGIN

## Query Match

## Best Local Similarity

## Matches

## Conservative

## Mismatches

## Indels

## Gaps

## CDS

## ORIGIN

## Query Match

## Best Local Similarity

## Matches

## Conservative

## Mismatches

## Indels

## Gaps

## CDS

## ORIGIN

## Query Match

## Best Local Similarity

## Matches

## Conservative

## Mismatches

## Indels

## Gaps

## CDS

## ORIGIN

## Query Match

## Best Local Similarity

## Matches

## Conservative

## Mismatches

## Indels

## Gaps

## CDS

## ORIGIN

## Query Match

## Best Local Similarity

## Matches

## Conservative

## Mismatches

## Indels

## Gaps

## CDS

## ORIGIN

## Query Match

## Best Local Similarity

## Matches

## Conservative

## Mismatches

## Indels

## Gaps

## CDS

## ORIGIN

```

QY      522 ACCGCCAAGTCTGAAATCTGGGCGCTGATTCTGGGCATCCTCATGACCATTTGGATTC 581
DB      241 ACCGCCAAGTCTGAAATCTGGGCGCTGATTCTGGGCATCCTCATGACCATTTGGATTC 300
QY      582 ATTCCTGTTACTGTTATTCGGCTCTGTGACAGTCTACCATATTTATTTACAGATATATACAG 641
DB      301 ATTCCTGTTACTGTTATTCGGCTCTGTGACAGTCTACCATATTTATTTACAGATATATACAG 360
QY      642 GAAACACGGGGTTACTAG 659
DB      361 GAAACACGGGGTTACTTG 378

RESULT 12
LOCUS   AP006283/c
DEFINITION Homo sapiens chromosome 11 clone RP11-326C3 map 11p, WORKING DRAFT
ACCESSION AP006283
VERSION   1
KEYWORDS HTG; HTGS PHASE1; HTGS_DRAFT.
SOURCE   Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE
AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE    Homo sapiens genomic DNA of 11p
JOURNAL  Published Only in Database (2003)
REFERENCE
AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE    Direct Submission
JOURNAL  Submitted (24-MAR-2003) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:hattori@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
----- Genomic Center
Center: RIKEN Genomic Sciences Center (GSC)
Center code: RIKEN
Web site: http://hgp.gsc.riken.go.jp/
Contact: hattori@gsc.riken.go.jp
----- Project Information
Center project name: HumDraIt11
Center clone name: RP11-326C3
----- Summary Statistics
Sequencing vector: PCR products; 100% of reads
Chemistry: Dye-terminator En-amersham; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 144706 bases at least Q40
Consensus quality: 150967 bases at least Q30
Consensus quality: 153657 bases at least Q20
Insert size: 155019; sum-of-ctngs
Quality coverage: 6.36x in Q20 bases; sum-of-ctngs
-----
NOTE: This is a 'working draft' sequence. It currently consists of
19 contigs. The true order of the pieces is not known and their
order in this sequence record is arbitrary. Gaps between the
contigs are represented as runs N, but the exact sizes of the gaps
are unknown. This record will be updated with the finished sequence
as soon as it is available and the accession number will be
preserved
1      21001 contig of 21001 bp in length
21102 contig of 19613 bp in length
40815 contig of 17616 bp in length
58430 contig of 14280 bp in length
66855 contig of 8529 bp in length
81235 contig of 12026 bp in length
89864 contig of 10636 bp in length
101990 contig of 10636 bp in length
112625 contig of 10636 bp in length
119802 contig of 7077 bp in length
119903 contig of 5279 bp in length
125182 contig of 100 bp
125184 contig of 3903 bp in length
129185 contig of 100 bp
129185 contig of 7976 bp in length
129285 contig of 100 bp
137261 contig of 6182 bp in length
137361 contig of 100 bp
143543 gap of 100 bp
143642 gap of 100 bp
143643 gap of 100 bp
146891 contig of 3249 bp in length
146892 contig of 100 bp
146992 contig of 2577 bp in length
149569 contig of 100 bp
149569 contig of 3010 bp in length
149569 contig of 100 bp
152679 gap of 100 bp
152779 gap of 100 bp
154135 contig of 1357 bp in length
154236 gap of 100 bp
154236 gap of 1427 bp in length
154236 contig of 100 bp
155663 gap of 100 bp
155763 contig of 1057 bp in length.

FEATURES
Location/Qualifiers
1..156819
/organism="Homo sapiens"
/mol_type="Genomic DNA"
/db_xref="taxon:9606"
/chromosome="11"
/map="11p"
/clone="RP11-326C3"
1..21001
/feature="assembly_fragment"
21102..40714
/feature="assembly_fragment"
40815..58430
/feature="assembly_fragment"
58531..66754
/feature="assembly_fragment"
66855..81134
/feature="assembly_fragment clone_end:SP6 vector_side:left"
misc_feature
misc_feature
misc_feature
misc_feature
misc_feature

```

```

/misc_feature      /note="assembly_fragment"
81235..89763
/misc_feature      /note="assembly_fragment"
89864..101889
/misc_feature      /note="assembly_fragment"
101900..112625
/misc_feature      /note="assembly_fragment"
112726..119802
/misc_feature      /note="assembly_fragment"
119903..125181
/misc_feature      /note="assembly_fragment"
125282..129184
/misc_feature      /note="assembly_fragment clone_end:T7 vector_side:right"
129285..137260
/misc_feature      /note="assembly_fragment"
137361..143542
/misc_feature      /note="assembly_fragment"
143643..146891
/misc_feature      /note="assembly_fragment"
146992..149568
/misc_feature      /note="assembly_fragment"
149669..152678
/misc_feature      /note="assembly_fragment"
152779..154135
/misc_feature      /note="assembly_fragment"
154236..155662
/misc_feature      /note="assembly_fragment"
155763..156813
/misc_feature      /note="assembly_fragment"

ORIGIN
Query Match      32.6%; Score 362; DB 2; Length 156819;
Best Local Similarity 100.0%; Pred. No. 1.1e-75;
Matches 362; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 106 AGCAGGAATAGAACTTAAGAGAAATACACACTTCTGAGAACTGAAACGACAGGGAA 165
DB 78353 AGCAGGAATAGAACTTAAGAGAAATACACACTTCTGAGAACTGAAACGACAGGGAA 78294
QY 166 AGGAGGTCTACTGAGCAGCGTCCAGACATCCGGACACACAGCGGCCCTTCCTCCACG 225
DB 78293 AGGAGGTCTACTGAGCAGCGTCCAGACATCCGGACACACAGCGGCCCTTCCTCCACG 78234
QY 226 CAGAAACACACACTTCTCAAACTTCACTCAACACTTCTTCCCAAGACGAGATGC 285
DB 78233 CAGAAACACACACTTCTCAAACTTCACTCAACACTTCTTCCCAAGACGAGATGC 78174
QY 286 ACAAGGAGGAAATGAGTGGTGTCTGGGGGACCCCGGACACATCTTCCCAAGGT 345
DB 78173 ACAAGGAGGAAATGAGTGGTGTCTGGGGGACCCCGGACACATCTTCCCAAGGT 78114
QY 346 CCACCGTGATCAACATCCACAGCAGACCTCCGTGCGCCGACCATGTGCTGTCCTGT 405
DB 78113 CCACCGTGATCAACATCCACAGCAGACCTCCGTGCGCCGACCATGTGCTGTCCTGT 78054
QY 406 TCACACACCTCTTCTGAACTGCTGTCTGGGCTTATAGCATTCGCTACTCCGTGA 465
DB 78053 TCACACACCTCTTCTGAACTGCTGTCTGGGCTTATAGCATTCGCTACTCCGTGA 77994
QY 466 AG 467
DB 77993 AG 77992

RESULT 13
AP006286/c      175416 bp      DNA      linear      HTG 26-MAR-2003
LOCUS
DEFINITION      Homo sapiens chromosome 11 clone RP13-25N22 map 11p, WORKING DRAFT
SEQUENCE, 22 unordered pieces.
ACCESSION      AP006286
VERSION
KEYWORDS      HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Homo sapiens (human)

```

ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1  
Hattori, M., Ishii, K., Toyoda, A., Taylor, T. D., Hong-Seog, P.,  
Fujiyama, A., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.  
Homo sapiens genomic DNA of 11p  
Published Only in Database (2003)  
2  
(bases 1 to 175416)  
Hattori, M., Ishii, K., Toyoda, A., Taylor, T. D., Hong-Seog, P.,  
Fujiyama, A., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.  
Direct Submission  
Submitted (24-MAR-2003) Masahira Hattori, The Institute of Physical  
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);  
1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
(E-mail: hattori@gsr.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/,  
Tel: 81-45-503-9111, Fax: 81-45-503-9170)  
----- Genome Center  
Center: RIKEN Genomic Sciences Center (GSC)  
Center code: RIKEN  
Web site: http://hgp.gsc.riken.go.jp/  
Contact: hattori@gsr.riken.go.jp  
----- Project Information  
Center project name: HumDraft11  
Center clone name: RP13-25N22  
----- Summary Statistics  
Sequencing vector: PCR products; 100% of reads  
Chemistry: Dye-terminator ET-amersham; 100% of reads  
Assembly program: Phrap; version 0.990329  
Consensus quality: 160074 bases at least Q40  
Consensus quality: 171662 bases at least Q20  
Insert size: 173316; sum-of-contigs  
Quality coverage: 7.92x in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of 22 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved

1 33131 contig of 33131 bp in length  
33232 54098 contig of 28667 bp in length  
54199 72037 contig of 17839 bp in length  
72138 83277 contig of 11140 bp in length  
83378 98599 contig of 15222 bp in length  
98700 105513 contig of 6814 bp in length  
105614 116229 contig of 10616 bp in length  
116330 121964 contig of 5635 bp in length  
122065 127627 contig of 5563 bp in length  
127728 131971 contig of 4244 bp in length  
132072 137506 contig of 5435 bp in length  
137607 143829 contig of 6223 bp in length  
143930 148539 contig of 4610 bp in length  
148640 153148 contig of 4509 bp in length  
153249 157509 contig of 4261 bp in length  
157610 160561 contig of 2952 bp in length  
160662 164778 contig of 4117 bp in length  
164879 168269 contig of 3391 bp in length  
168370 170689 contig of 2320 bp in length  
170790 172833 contig of 2044 bp in length  
172934 174155 contig of 1222 bp in length  
174256 175416 contig of 1161 bp in length.

\* NOTE: This is a 'working draft' sequence. It currently consists of 22 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.  
1 33131: contig of 33131 bp in length

33132 33231: gap of 100 bp  
 33232 54098: contig of 20867 bp in length  
 54099 54198: gap of 100 bp  
 54199 72037: contig of 17839 bp in length  
 72038 83277: gap of 100 bp  
 83278 83277: contig of 11140 bp in length  
 83278 98599: gap of 100 bp  
 98600 98699: contig of 15222 bp in length  
 98700 105513: gap of 100 bp  
 105514 105614: gap of 100 bp  
 105615 116229: contig of 10616 bp in length  
 116230 116230: gap of 100 bp  
 116231 121964: contig of 5635 bp in length  
 121965 122064: gap of 100 bp  
 122065 127627: contig of 5563 bp in length  
 127628 127727: gap of 100 bp  
 127728 131971: contig of 4244 bp in length  
 131972 132071: gap of 100 bp  
 132072 137506: contig of 5435 bp in length  
 137507 137606: gap of 100 bp  
 137607 143829: contig of 6223 bp in length  
 143830 143929: gap of 100 bp  
 143930 148539: contig of 4610 bp in length  
 148540 148639: gap of 100 bp  
 148640 153148: contig of 4509 bp in length  
 153149 153248: gap of 100 bp  
 153249 157509: contig of 4261 bp in length  
 157510 157609: gap of 100 bp  
 157610 160561: contig of 2952 bp in length  
 160562 160661: gap of 100 bp  
 160662 164778: contig of 4117 bp in length  
 164779 164878: gap of 100 bp  
 164879 168269: contig of 3391 bp in length  
 168270 168369: gap of 100 bp  
 168370 170689: contig of 2320 bp in length  
 170690 170789: gap of 100 bp  
 170790 172833: contig of 2044 bp in length  
 172834 172933: gap of 100 bp  
 172934 174155: contig of 1222 bp in length  
 174156 174256: gap of 100 bp  
 174257 175416: contig of 1161 bp in length.

FEATURES

Location/Qualifiers  
 1..175416  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="11"  
 /map="11p"  
 /clone="RP13-25N22"

misc\_feature 1..33131  
 /note="assembly\_fragment clone\_end:SP6 vector\_side:left"  
 misc\_feature 33232..54098  
 /note="assembly\_fragment"  
 misc\_feature 54199..72037  
 /note="assembly\_fragment"  
 misc\_feature 72138..83277  
 /note="assembly\_fragment"  
 misc\_feature 83378..98599  
 /note="assembly\_fragment"  
 misc\_feature 98700..105513  
 /note="assembly\_fragment clone\_end:T7 vector\_side:left"  
 misc\_feature 105614..116229  
 /note="assembly\_fragment"  
 misc\_feature 116330..121964  
 /note="assembly\_fragment"  
 misc\_feature 122065..127627  
 /note="assembly\_fragment"  
 misc\_feature 127728..131971  
 /note="assembly\_fragment"  
 misc\_feature 132072..137506  
 /note="assembly\_fragment"  
 misc\_feature 137607..143829

misc\_feature /note="assembly\_fragment"  
 143930..148539  
 /note="assembly\_fragment"  
 148640..153148  
 /note="assembly\_fragment"  
 153249..157509  
 /note="assembly\_fragment"  
 157610..160561  
 /note="assembly\_fragment"  
 160662..164778  
 /note="assembly\_fragment"  
 164879..168269  
 /note="assembly\_fragment"  
 168370..170689  
 /note="assembly\_fragment"  
 170790..172833  
 /note="assembly\_fragment"  
 172934..174155  
 /note="assembly\_fragment"  
 174256..175416  
 /note="assembly\_fragment"

ORIGIN

Query Match 32.6%; Score 362; DB 2; Length 175416;  
 Best Local Similarity 100.0%; Pred.No. 1.1e-75;  
 Matches 362; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 106 AGCAGGAAATAGAACTTAAGAGAAATACACACTTCTGAGAAATGAAACGACAGGGGAA 165  
 DB 76527 AGCAGGAAATAGAACTTAAGAGAAATACACACTTCTGAGAAATGAAACGACAGGGGAA 76468  
 QY 166 AGGAGGTCCTCACTGAGCAGCGTCCAGCATCCGAGACACACAGCGGCCCTTCGCTCCACG 225  
 DB 76467 AGGAGGTCCTCACTGAGCAGCGTCCAGCATCCGAGACACACAGCGGCCCTTCGCTCCACG 76408  
 QY 226 CAGAAACACACACTTCTCAACCTTCACTCAACACTTCTTCCCAAGCCAGAGATGC 285  
 DB 76407 CAGAAACACACACTTCTCAACCTTCACTCAACACTTCTTCCCAAGCCAGAGATGC 76348  
 QY 286 ACAAGGAGGAACATGAGTGCTGTGTGGGGACACCCCGGACACCATCTCTTCAAGGT 345  
 DB 76347 ACAAGGAGGAACATGAGTGCTGTGTGGGGACACCCCGGACACCATCTCTTCAAGGT 76288  
 QY 346 CCACCGTGATCAACATCCACAGCAGACCTCGTGCCCGACCATGCTGCTGCTGCTGCTGT 405  
 DB 76287 CCACCGTGATCAACATCCACAGCAGACCTCGTGCCCGACCATGCTGCTGCTGCTGCTGT 76228  
 QY 406 TCAACACCTCTTCTTGAACCTGTGCTGTGGGCTTCATAGCATTCGCTTACTCGGTGA 465  
 DB 76227 TCAACACCTCTTCTTGAACCTGTGCTGTGGGCTTCATAGCATTCGCTTACTCGGTGA 76168  
 QY 466 AG 467  
 DB 76167 AG 76166

RESULT 14

AC138230 187160 bp DNA linear PRI 22-MAY-2003  
 LOCUS Homo sapiens chromosome 11, clone RP13-317D12, complete sequence.  
 DEFINITION AC138230  
 ACCESSION AC138230  
 VERSION AC138230.5 GI:30984765  
 KEYWORDS HTG.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 187160)  
 AUTHORS Birren,B., Nusbaum,C. and Lander,E.  
 TITLE Homo sapiens chromosome 11, clone RP13-317D12  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 187160)  
 AUTHORS Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,



Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Boukhgalter,B., Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A., Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J., Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hafez,N., Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,G., Maclean,C., Macdonald,P., Major,J., Matthews,C., McCarthy,M., Meldrim,J., Meneus,L., Meneus,L., Meneus,L., Mhova,T., Mienga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Peterson,K., Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Riese,C., Rogov,P., Roy,A., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talanas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission  
 Submitted (20-DEC-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
 3 (bases 1 to 187160)

REFERENCE  
 AUTHORS  
 Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y., Collymore,A., Cook,A., Cooke,P., Corum,B., DeArellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,G., Lui,A., Mabbitt,R., Maclean,C., Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M., Meldrim,J., Meneus,L., Mihoa,T., Mienga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Rachupka,A., Ramasany,U., Raymond,C., Retta,R., Riese,C., Rogov,P., Roman,J., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission  
 Submitted (10-APR-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
 4 (bases 1 to 187160)

REFERENCE  
 AUTHORS  
 Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y., Collymore,A., Cook,A., Cooke,P., Corum,B., DeArellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,G., Lui,A., Mabbitt,R., Maclean,C., Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M., Meldrim,J., Meneus,L., Mihoa,T., Mienga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Rachupka,A., Ramasany,U., Raymond,C., Retta,R., Riese,C., Rogov,P., Roman,J., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission  
 Submitted (22-MAY-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
 On May 22, 2003 this sequence version replaced gi:29423936.  
 All repeats were identified using RepeatMasker:  
 Smit,A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research  
 Center code: WIBR  
 Web site: <http://www-seq.wi.mit.edu>  
 Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)  
 ----- Project Information  
 Center project name: L29013  
 Center clone name: 317\_D12  
 -----

FEATURES	Location/Qualifiers
source	1..187160
repeat_region	/organism="Homo sapiens"
repeat_region	/mol_type="genomic DNA"
repeat_region	/db_xref="taxon:9606"
repeat_region	/chromosome="11"
repeat_region	/map="11"
repeat_region	/clone="RP13-317D12"
repeat_region	/clone_lib="RPCI-13 Human Female BAC"
repeat_region	135..200
repeat_region	/rpt_family="L2"
repeat_region	1134..1310
repeat_region	/rpt_family="MIR"
repeat_region	complement(1311..1614)
repeat_region	/rpt_family="AluY"
repeat_region	1615..1638
repeat_region	/rpt_family="MIR"
repeat_region	1849..1938
repeat_region	/rpt_family="L1ME"
repeat_region	complement(4297..4385)
repeat_region	/rpt_family="L2"
repeat_region	4463..4561
repeat_region	/rpt_family="PEAM"
repeat_region	complement(4573..4613)
repeat_region	/rpt_family="MADE1"
repeat_region	4733..5040
repeat_region	/rpt_family="AluY"
repeat_region	complement(5094..5383)
repeat_region	/rpt_family="L2"
repeat_region	complement(5379..5505)
repeat_region	/rpt_family="L2"
repeat_region	complement(5501..5627)
repeat_region	/rpt_family="L2"
repeat_region	complement(5667..6096)
repeat_region	/rpt_family="L2"
repeat_region	6542..6666
repeat_region	/rpt_family="MER41E"
repeat_region	6699..6719
repeat_region	/rpt_family="GC-rich"
repeat_region	6917..6996
repeat_region	/rpt_family="C-rich"
repeat_region	7136..7329
repeat_region	/rpt_family="C-rich"
repeat_region	15461..15483
repeat_region	/rpt_family="(TG)n"
repeat_region	18178..18413
repeat_region	/rpt_family="L2"
repeat_region	complement(18543..18907)
repeat_region	/rpt_family="MER4A"
repeat_region	complement(18913..19138)
repeat_region	/rpt_family="AluSg/x"
repeat_region	complement(19139..19318)
repeat_region	/rpt_family="MER4D"
repeat_region	complement(19319..19688)
repeat_region	/rpt_family="MER4E1"
repeat_region	19689..19992
repeat_region	/rpt_family="AluSp"
repeat_region	19994..20308
repeat_region	/rpt_family="AluSg"
repeat_region	complement(20309..20630)
repeat_region	/rpt_family="MER4E1"
repeat_region	20632..20763
repeat_region	/rpt_family="AluSg/x"
repeat_region	complement(20785..20882)
repeat_region	/rpt_family="MER4E"

repeat_region	complement(20886..21091)	AC136475	143779 bp	DNA	linear	PRI 08-JUL-2003
repeat_region	/rpt_family="MER41E"	AC136475				
repeat_region	/rpt_family="AluX"	AC136475				
repeat_region	complement(21403..21580)	AC136475				
repeat_region	/rpt_family="MER41E"	AC136475				
repeat_region	complement(22261..22538)	AC136475				
repeat_region	/rpt_family="MER41E"	AC136475				
repeat_region	complement(22543..22594)	AC136475				
repeat_region	/rpt_family="MER4D"	AC136475				
repeat_region	22595..22902	AC136475				
repeat_region	/rpt_family="AluY"	AC136475				
repeat_region	complement(22903..23195)	AC136475				
repeat_region	/rpt_family="MER4D"	AC136475				
repeat_region	23197..23314	AC136475				
repeat_region	/rpt_family="FLAM_C"	AC136475				
repeat_region	complement(23318..23368)	AC136475				
repeat_region	/rpt_family="MER4D"	AC136475				
repeat_region	23369..23890	AC136475				
repeat_region	/rpt_family="MER4A"	AC136475				
repeat_region	complement(23891..24166)	AC136475				
repeat_region	/rpt_family="MER4D"	AC136475				
repeat_region	complement(24381..24698)	AC136475				
repeat_region	/rpt_family="MER63"	AC136475				
repeat_region	complement(25372..25668)	AC136475				
repeat_region	/rpt_family="AluY"	AC136475				
repeat_region	26402..26493	AC136475				
repeat_region	/rpt_family="(TG)n"	AC136475				
repeat_region	complement(27325..28044)	AC136475				
repeat_region	/rpt_family="LTR30"	AC136475				
repeat_region	28747..28755	AC136475				
repeat_region	/notes"<30 qual SNGL region"	AC136475				
repeat_region	29024..29047	AC136475				
repeat_region	/rpt_family="GC rich"	AC136475				
repeat_region	complement(29397..30103)	AC136475				
repeat_region	/rpt_family="LTR8"	AC136475				
repeat_region	30337..30656	AC136475				
repeat_region	/rpt_family="LTR2C"	AC136475				
repeat_region	complement(30657..31083)	AC136475				
Query Match	32.6%; Score 362; DB 9; Length 187160;					
Best Local Similarity	100.0%; Pred. No. 1.1e-75;					
Matches 362; Conservative	0; Mismatches 0; Indels 0; Gaps 0;					
QY	106 AGCAGGAATAGAACTTACAGAGAAATACACATCTTCTGAGAACTGAAACGACAGGGAA 165					
Db	31874 AGCAGGAATAGAACTTACAGAGAAATACACATCTTCTGAGAACTGAAACGACAGGGAA 31933					
QY	166 AGGAGTCTCAGTACGACCGTCCAGCATCGGACACACAGCGGCCCTTCGCTCCACG 225					
Db	31934 AGGAGTCTCAGTACGACCGTCCAGCATCGGACACACAGCGGCCCTTCGCTCCACG 31993					
QY	226 CAGAAACACACATCTTCAACCTTCACTCAACATCTTCTTCCCAAGCCAGAGATGC 285					
Db	31994 CAGAAACACACATCTTCAACCTTCACTCAACATCTTCTTCCCAAGCCAGAGATGC 32053					
QY	286 ACAGGAGGACATGAGTGGCTGTGTGGGGGACACCCCGACACCATCTTCCCAAGT 345					
Db	32054 ACAGGAGGACATGAGTGGCTGTGTGGGGGACACCCCGACACCATCTTCCCAAGT 32113					
QY	346 CCACCGTGATCAACATCCACAGCAGACCTTCGTCGCCGACCATGTGCTGCTCCCTGT 405					
Db	32114 CCACCGTGATCAACATCCACAGCAGACCTTCGTCGCCGACCATGTGCTGCTCCCTGT 32173					
QY	406 TCAACACCTCTTTGAACTGGTCTGTCTGGGGTTCTAGCATTCGCTACTCGCTCA 465					
Db	32174 TCAACACCTCTTTGAACTGGTCTGTCTGGGGTTCTAGCATTCGCTACTCGCTCA 32233					
QY	466 AG 467					
Db	32234 AG 32235					

RESULT 15	AC136475
LOCUS	AC136475
DEFINITION	Homo sapiens chromosome 11, clone RP11-326C3, complete sequence.
ACCESSION	AC136475
VERSION	AC136475.7 GI:32469539
KEYWORDS	HTG.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	Birren,B., Nusbaum,C. and Lander,E.
TITLE	1 (bases 1 to 143779)
JOURNAL	Homo sapiens chromosome 11, clone RP11-326C3
REFERENCE	Unpublished
AUTHORS	2 (bases 1 to 143779)
REFERENCE	Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Boukhgalter,B., Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A., Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J., Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hafez,N., Hagoian,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,G., Maclean,C., Macdonald,P., Major,J., Matthews,C., McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Minova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P., Roman,J., Roy,A., Schauer,S., Schuback,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
TITLE	Direct Submission
JOURNAL	Submitted (02-NOV-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE	3 (bases 1 to 143779)
AUTHORS	Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y., Collymore,A., Cook,A., Cooke,P., Corum,B., DeArellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N., Hagoian,J., Hagoian,B., Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,G., Lui,A., Mabbitt,R., Maclean,C., Macdonald,P., Major,J., Manning,J., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schuback,R., Seaman,S., Severy,P., Smith,C., Stange-Thomann,N., Stojanovic,N., Talamas,J., Tesfaye,S., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
TITLE	Direct Submission
JOURNAL	Submitted (10-APR-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE	4 (bases 1 to 143779)
AUTHORS	Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y., Collymore,A., Cook,A., Cooke,P., Corum,B., DeArellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N., Hagoian,J., Hagoian,B., Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,G., Lui,A., Mabbitt,R., Maclean,C., Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,

Meldrim, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schupback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talanas, J., Testaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V. S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

# **TITLE**

**JOURNAL** Submitted (30-MAY-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

## **REFERENCE**

## **AUTHORS**

5 (bases 1 to 143779)  
 Birren, B., Nusbaum, C., Lander, E., Aboueleil, A., Allen, N., Anderson, M., Arachchi, H. M., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhgalter, B., Camarata, J., Chang, J., Choepel, Y., Collymore, A., Cook, A., Cooke, P., Corum, B., DeArelano, K., Diaz, J. S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S., Ferreira, P., FitzGerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kellis, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, X., Lui, A., Mabbitt, R., MacLean, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schupback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talanas, J., Testaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V. S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

**Direct Submission**  
 Submitted (08-JUN-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
 On Jul 8, 2003 this sequence version replaced gi:31193978.  
 All repeats were identified using RepeatMasker:  
 Smith, A. F. A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

## **TITLE**

## **JOURNAL**

## **COMMENT**

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)

----- Project Information

Center project name: L28535

Center clone name: 326\_C\_3

-----

Only the first 143.8 kilobases of this clone are being submitted.  
 The remainder overlaps accession number AC138230 [WICGR project L29013].

## **FEATURES**

## **source**

Location/Qualifiers  
 1..143779  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="11"  
 /map="11"  
 /clone="RP11-326C3"  
 /clone\_lib="RP11 Human Male BAC"  
 1712..2006  
 /rpt\_family="AluSx"

## **repeat\_region**

## **repeat\_region**

## **repeat\_region**

## **repeat\_region**

## **repeat\_region**

## **repeat\_region**

## **repeat\_region**

## **repeat\_region**

## **repeat\_region**

## **repeat\_region**

## **repeat\_region**

repeat\_region  
 4902..4968  
 /rpt\_family="AluSx"  
 repeat\_region  
 4979..5091  
 /rpt\_family="AT\_rich"  
 repeat\_region  
 5154..5251  
 /rpt\_family="L1MB4a"  
 repeat\_region  
 5620..5683  
 /rpt\_family="L1MB4a"  
 repeat\_region  
 complement(5684..5954)  
 /rpt\_family="AluJo"  
 repeat\_region  
 5955..6050  
 /rpt\_family="L1MB4a"  
 repeat\_region  
 complement(6059..6206)  
 /rpt\_family="MER46C"  
 repeat\_region  
 6349..6578  
 /rpt\_family="L1MB4a"  
 repeat\_region  
 6792..7203  
 /rpt\_family="L2"  
 repeat\_region  
 complement(884..9052)  
 /rpt\_family="L1MC5"  
 repeat\_region  
 9225..9287  
 /rpt\_family="MER68-int"  
 repeat\_region  
 9676..9986  
 /rpt\_family="AluSg"  
 repeat\_region  
 9994..10032  
 /rpt\_family="AT\_rich"  
 repeat\_region  
 10182..10417  
 /rpt\_family="L1MB4a"  
 repeat\_region  
 complement(10418..10946)  
 /rpt\_family="L1MB4a"  
 repeat\_region  
 complement(10947..11251)  
 /rpt\_family="AluSx"  
 repeat\_region  
 complement(11252..11272)  
 /rpt\_family="L1MB4a"  
 repeat\_region  
 complement(11290..11595)  
 /rpt\_family="AluSg"  
 repeat\_region  
 11702..12014  
 /rpt\_family="AluSg"  
 repeat\_region  
 12047..12137  
 /rpt\_family="L1M4"  
 repeat\_region  
 complement(12159..12465)  
 /rpt\_family="AluSx"  
 repeat\_region  
 12466..12805  
 /rpt\_family="L2"  
 repeat\_region  
 12763..12832  
 /rpt\_family="L2"  
 repeat\_region  
 12898..13715  
 /rpt\_family="L1MB5"  
 repeat\_region  
 13716..14007  
 /rpt\_family="AluSg"  
 repeat\_region  
 14018..14320  
 /rpt\_family="AluSx"  
 repeat\_region  
 14321..14603  
 /rpt\_family="L1MB5"  
 repeat\_region  
 14611..14772

Query Match 32.5%; Score 360.4; DB 9; Length 143779;

Best Local Similarity 99.7%; Pred. No. 2.7e-75;

Matches 361; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 106 AGCAGGAATAAGAACTTAAGAGAAATACACACTTCTGAGAACTGAAACGACAGGGGAA 165  
 DB 98539 AGCAGGAATAAGAACTTAAGAGAAATACACACTTCTGAGAACTGAAACGACAGGGGAA 98598  
 QY 166 AGGAGGTCTCTAGCAGCCGTCGCCAGCATCCGACACACAGGGGCTTCGCTCCACG 225  
 DB 98599 AGGAGGTCTCTAGCAGCCGTCGCCAGCATCCGACACACAGGGGCTTCGCTCCACG 98658  
 QY 236 CAGAAAACACACTTCTCAAACTTCACTCAACACTTCTTCCCAAGCAGAGATGC 285  
 DB 98659 CAGAAAACACACTTCTCAAACTTCACTCAACACTTCTTCCCAAGCAGAGATGC 98718

QY	286	ACAAGGAGGAACATGAGGTGGCTGTGCTGGGGGACCCCCCAGCACCACATCCTTCCAAGGT	345
Db	98719	ACAAGGAGGAACATGAGGTGGCTGTGCTGGGGGACCCCCCAGCACCACATCCTTCCAAGGT	98778
QY	346	CCACCGTGATCAACATCCACAGCGAGACCTCCGTGCCCGACCATGTGCTGGTCCCTGT	405
Db	98779	CCACCGTGATCAACATCCACAGCGAGACCTCCGTGCCCGACCATGTGCTGGTCCCTGT	98838
QY	406	TCAACACCCCTCTTCTTGAACCTGGTGTGCTGTGGGCTTCATAGCATTCGCCTACTCCGTGA	465
Db	98839	TCAACACCCCTCTTCTTGAACCTGGTGTGCTGTGGGCTTCATAGCATTCGCCTACTCCGTGA	98898
QY	466	AG 467	
Db	98899	AG 98900	

Search completed: March 11, 2004, 03:36:53  
 Job time : 2828.74 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 10, 2004, 08:01:57 ; Search time 269.909 Seconds  
(without alignments)  
17454.974 Million cell updates/sec

Title: US-10-084-817-37  
Perfect score: 1109  
Sequence: 1 cctgcacaggagacactgg.....acanaattccaccctatccg 1109

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_29Jan04:\*  
1: Geneseqn1980s:\*  
2: Geneseqn1990s:\*  
3: Geneseqn2000s:\*  
4: Geneseqn2001as:\*  
5: Geneseqn2001bs:\*  
6: Geneseqn2002s:\*  
7: Geneseqn2003as:\*  
8: Geneseqn2003bs:\*  
9: Geneseqn2003cs:\*  
10: Geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1051	94.8	1109	8	ACD42205 Human int
2	802	72.3	851	2	AAZ41369 Human nor
3	703	63.4	1617	9	ADE25685 Human CDN
4	699.8	63.1	1095	6	AAS95008 Human DNA
5	696.2	62.8	842	6	ABK46192 CDNA enco
6	696.2	62.8	853	5	ABA83110 IFN-induc
7	696.2	62.8	853	6	ABK83770 Human CDN
8	696.2	62.8	853	6	ABK64497 Human ben
9	696.2	62.8	853	6	ABN95610 Gene #210
10	685.8	61.8	852	7	ABT31907 Human bre
11	624.4	56.3	647	6	ABQ60782 Human 9-2
12	624.4	56.3	647	7	ABX10351 DNA enco
13	624.4	56.3	647	9	ADD18965 Human dis
14	562	50.7	643	6	ABQ60669 Human col
15	553.4	49.9	651	9	ADE07174 Novel cod
16	552.8	49.8	674	6	ABQ60529 Human col
17	541.4	48.8	619	6	ABQ60439 Human col
18	518.8	46.8	579	6	ABQ60650 Human col
19	511	46.1	640	6	ABQ59690 Human col
20	481.6	43.4	527	6	ABQ60421 Human col
21	480	43.3	581	6	ABQ60275 Human col
22	464.4	41.9	676	6	ABQ60312 Human col
23	454.8	41.0	701	6	ABQ55533 Human ova

C	24	441.2	39.8	519	6	ABQ60383	Human col
	25	435	39.2	689	6	ABQ60308	Human col
	26	433	39.0	462	8	ACH32390	Human end
	27	427.6	38.6	434	6	ACH48424	Human leu
C	28	390.6	35.2	593	6	ABQ60272	Human col
	29	376.4	33.9	378	6	ABSS1531	Human CDN
	30	373.4	33.7	376	5	AAS77261	DNA enco
	31	373.4	33.7	376	5	AAS77263	DNA enco
	32	373.4	33.7	376	9	ADE08931	Novel DNA
	33	364	32.8	384	5	AAS93007	DNA enco
	34	360.8	32.5	388	5	AAS75775	DNA enco
	35	355.8	32.1	598	6	ABQ60320	Human col
C	36	354	31.9	911	5	AAS75528	Human col
	37	353.2	31.8	507	4	AAH35435	Human col
	38	351	31.7	352	6	ABK39166	cDNA enco
	39	351	31.7	352	7	ACA11495	Human lun
	40	351	31.7	352	7	ACR02681	Lung canc
	41	348.2	31.4	400	2	AAT22123	Human gen
	42	340.8	30.7	622	9	ADE81113	Bovine le
	43	335.8	30.3	764	2	AAV69033	DNA molec
C	44	335.8	30.3	764	3	AAC80995	Human bre
C	45	335.8	30.3	764	6	AAS99841	Breast tu

ALIGNMENTS

RESULT 1  
ID ACD42205  
AC ACD42205 standard; cDNA; 1109 BP.  
AC ACD42205;  
DT 05-SEP-2003 (first entry)  
DE Human interferon inducible protein cDNA Incyte 3094768CB1.  
KW Human; ss; DNA methylation; cancer; colon cancer.  
OS Homo sapiens.  
XX  
XX US2003013099-A1.  
XX  
XX 16-JAN-2003.  
XX  
XX 07-MAR-2002; 2002US-00093766.  
XX  
XX 19-MAR-2001; 2001US-0277380P.  
PA (LASEK/) LASEK A K W.  
PA (JONES/) JONES D A.  
PA (KARP/) KARP A R.  
PI Lasek AKW, Jones DA, Karpf AR;  
XX  
XX WPI; 2003-503249/47.  
XX P-PSDB; ABO25116.  
XX  
XX New combination comprising cDNAs that are expressed in a disorder or  
XX process associated with DNA methylation, useful for diagnosing, staging,  
XX treating or monitoring treatment of cancer, e.g. colon cancer.  
XX  
XX Claim 2; Page 22; 66pp; English.  
XX  
XX The invention relates to a combination comprising cDNAs which are  
XX expressed in a disorder or process associated with DNA methylation. The  
XX combination and cDNAs are useful for diagnosing, staging, treating or  
XX monitoring treatment of cancer, e.g. colon cancer and for detecting or  
XX changes in expression of genes encoding proteins that are associated with  
XX DNA methylation. The protein is useful for screening molecules or  
XX compounds to identify at least one ligand that binds to the protein and  
XX for producing an antibody. The present sequence represents a cDNA  
XX expressed in a disorder or process associated with DNA methylation

```

XX SQ Sequence 1109 BP; 268 A; 340 C; 241 G; 202 T; 0 U; 58 Other;
Query Match 94.8%; Score 1051; DB 8; Length 1109;
Best Local Similarity 100.0%; Pred. No. 3.3e-277;
Matches 1109; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCTGCACGAGAGACACTGGGAGGTTAGTCCCAAAACCGCGACAGAGCAGGACTGCAGC 60
Db 1 CCTGCACGAGAGACACTGGGAGGTTAGTCCCAAAACCGCGACAGAGCAGGACTGCAGC 60

Qy 61 CTGAGAAAGACCAAGATTCTCAGGAGAGAGGCTGCGACAAGTGACAGCAAGAAATAGAAA 120
Db 61 CTGAGAAAGACCAAGATTCTCAGGAGAGAGGCTGCGACAAGTGACAGCAAGAAATAGAAA 120

Qy 121 CTTAAGAGAAATACACACTTCTGAGAAAATGAAACGACAGGGGAAAGAGGTCTCACTGA 180
Db 121 CTTAAGAGAAATACACACTTCTGAGAAAATGAAACGACAGGGGAAAGAGGTCTCACTGA 180

Qy 181 GCACGCTCCAGCATCCGACACACACAGCGGCTTCTGCTCCAGCGAGAAAACCACTT 240
Db 181 GCACGCTCCAGCATCCGACACACACAGCGGCTTCTGCTCCAGCGAGAAAACCACTT 240

Qy 241 CTCAAACTTCACTCAACACTTCTTCCCAAGCCAGAGATGCAAGAGGAAACATG 300
Db 241 CTCAAACTTCACTCAACACTTCTTCCCAAGCCAGAGATGCAAGAGGAAACATG 300

Qy 301 AGGTGGCTGTGCTGGGGGACCCCGACACCATCTTCCCAAGGTCACCGTGATCAACA 360
Db 301 AGGTGGCTGTGCTGGGGGACCCCGACACCATCTTCCCAAGGTCACCGTGATCAACA 360

Qy 361 TCCACAGGAGAGACCTCGCTGCGCGACCATGCTCTGTGCTCTTCAACACCTCTTCT 420
Db 361 TCCACAGGAGAGACCTCGCTGCGCGACCATGCTCTGTGCTCTTCAACACCTCTTCT 420

Qy 421 TGAACCTGTGCTGTGGGCTTCAATAGCATTCGCTACTCGTGAAGTCTAGGACAGGA 480
Db 421 TGAACCTGTGCTGTGGGCTTCAATAGCATTCGCTACTCGTGAAGTCTAGGACAGGA 480

Qy 481 AGATGGTTGGGACGTCAGCGGGGCGCCGATGCTCCACCGCAAGTGCCTGAACA 540
Db 481 AGATGGTTGGGACGTCAGCGGGGCGCCGATGCTCCACCGCAAGTGCCTGAACA 540

Qy 541 TCTGGGCTGTGCTGGGACATCTATGACCATTTGGATTCTCTTACTGTATTCG 600
Db 541 TCTGGGCTGTGCTGGGACATCTATGACCATTTGGATTCTCTTACTGTATTCG 600

Qy 601 GCTCTGTGACAGTCTACCATATTATGTTACAGATAATACAGGAAAACGGGTTACTAGT 660
Db 601 GCTCTGTGACAGTCTACCATATTATGTTACAGATAATACAGGAAAACGGGTTACTAGT 660

Qy 661 AGCGGCCCATAGCTGCAACCTTTCGACTCCACTGTGCAATGCTGCGCTGACGCTGG 720
Db 661 AGCGGCCCATAGCTGCAACCTTTCGACTCCACTGTGCAATGCTGCGCTGACGCTGG 720

Qy 721 GCTGTGCTCCCTGCCCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
Db 721 GCTGTGCTCCCTGCCCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780

Qy 781 GTCTACACTGACATTCATTAAGTGACGCTGCTGTGTAAGAAAACCAATAAACCCGAG 840
Db 781 GTCTACACTGACATTCATTAAGTGACGCTGCTGTGTAAGAAAACCAATAAACCCGAG 840

Qy 841 GGGGGGGCGGACCATTTCCGCTAAGGGGAGGATATACATTTCCGGCGGTGTATAC 900
Db 841 GGGGGGGCGGACCATTTCCGCTAAGGGGAGGATATACATTTCCGGCGGTGTATAC 900

Qy 901 AGCGTGGGATGGGACACCTTGGGTATCCAAATTAACGCTGATCCTTCAGAGNCCCGGG 960
Db 901 AGCGTGGGATGGGACACCTTGGGTATCCAAATTAACGCTGATCCTTCAGAGNCCCGGG 960

Qy 961 ATNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1020

```

```

Db 961 ATNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1020
Qy 1021 CNCCACCCNCCNCCCACTCTCCCCACACACCCCAACCAACACCCAGCCTCA 1080
Db 1021 CNCCACCCNCCNCCCACTCTCTCCCCACACACCCCAACCAACACCCAGCCTCA 1080
Qy 1081 CCATCCCTCACANAATTCCACCTATCCG 1109
Db 1081 CCATCCCTCACANAATTCCACCTATCCG 1109

RESULT 2
ID AA241369 standard; cDNA; 851 BP.
XX AA241369;
XX 19-JAN-2000 (first entry)
XX Human normal uterus tissue derived cDNA 45.
XX Human; uterus; cancer; treatment; anticancer; cytostatic; gene therapy;
KW EST; expressed sequence tag; ss.
XX Homo sapiens.
OS DE19817946-A1.
PN 21-OCT-1999.
PD 17-APR-1998; 98DE-01017946.
PF 17-APR-1998; 98DE-01017946.
PR 17-APR-1998; 98DE-01017946.
XX (META-) METAGEN GES GENOMFORSCHUNG MBH.
XX Rosenthal A, Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E;
PI WPI; 1999-591956/51.
DR New nucleic acid sequences expressed in normal uterine tissues, and
PT derived polypeptides, for treatment of uterine cancer and identification
PT of therapeutic agents.
XX Claim 3; Page 108; 154pp; German.
XX This invention describes novel cDNA sequences (A) highly expressed in
CC normal uterine tissue which can have anticancer and cytostatic activity
CC and can be used for gene therapy. (A) are used (i) for recombinant
CC expression of polypeptides (B) and (ii) to isolate complete genes. (B)
CC are used (i) to identify agents suitable for treatment of uterine cancer;
CC (ii) directly for treating this form of cancer (including expression from
CC gene therapy vectors) and (iii) for generation of specific antibodies.
CC (A) are identified by assembling ESTs (expressed sequence tags) from a
CC particular tissue type before comparison of expression patterns. This
CC allows a significantly longer fragment of the gene to be revealed, so
CC should reduce the number of failures associated with the fact that ESTs
CC from different libraries may represent different parts of the same
CC unknown gene, distorting the estimated frequency of occurrence in a
CC particular tissue. AA241325-241385 represent the human uterine tissue
CC derived cDNA fragments of the invention which encode the protein
CC fragments represented in AAY59838-Y59892
XX Sequence 851 BP; 225 A; 248 C; 211 G; 167 T; 0 U; 0 Other;

Query Match 72.3%; Score 802; DB 2; Length 851;
Best Local Similarity 98.6%; Pred. No. 4.4e-209;
Matches 830; Conservative 0; Mismatches 10; Indels 2; Gaps 2;

Qy 1 CCTGCACGAGAGACACTGGGAGGTTAGTCCCAAAACCGCGACAGAGCAGGACTGCAGC 60
Db 2 CCTGCACGAGAGACACTGGGAGGTTAGTCCCAAAACCGCGACAGAGCAGGACTGCAGC 61

```



QY 61 CTCAGGAAAGAGCAAGATTTTCAGGAGAGAGCGCTCGACAAAGTCAGCAGGAAATAGAAA 120  
DB 62 CTGAGGAAAGAGCAAGATTTTCAGGAGAGAGCGCTCGACAAAGTCAGCAGGAAATAGAAA 121  
QY 121 CTTAAGAGAAATACACACTTCTGTAGAAACTGAAACGACAGGAGGAGAGGTCTCACTGA 180  
DB 122 CTTAAGAGAAATACACACTTCTGTAGAAACTGAAACGACAGGAGGAGAGGTCTCACTGA 180  
QY 181 GCACCGTCCAGCATCCGACACACACAGCGCGCTTGGCTCCAGCAGAAACCACTT 240  
DB 181 GCACCGTCCAGCATCCGACACACAGCGCGCTTGGCTCCAGCAGAAACCACTT 240  
QY 241 CTCACACCTTCACTCAACACTTCTTCCCAAGCCAGAAATGACAGGAGGAGCAATG 300  
DB 241 CTCACACCTTCACTCAACACTTCTTCCCAAGCCAGAAATGACAGGAGGAGCAATG 300  
QY 301 AGTGGTGTGTGGGGGACCCCGACACATCTCTTCAAGGTCCACCGTATCAACA 360  
DB 301 AGTGGTGTGTGGGGGACCCCGACACATCTCTTCAAGGTCCACCGTATCAACA 360  
QY 361 TCCACAGCAGAGCTCCGTGCGCCGACCATGCTGTCTGCTTCAACACCTTCTTCT 420  
DB 361 TCCACAGCAGAGCTCCGTGCGCCGACCATGCTGTCTGCTTCAACACCTTCTTCT 420  
QY 421 TGAACCTGTGTGTGGGCTTCATAGCATTCGCTACTCCGTGAAGTCTAGGACAGGA 480  
DB 421 TGAACCTGTGTGTGGGCTTCATAGCATTCGCTACTCCGTGAAGTCTAGGACAGGA 480  
QY 481 AGATGTTGGCGAGTGCACCGGGGCCCGAGGCTATGCTCCACCGCAAGTCCCTGAACA 540  
DB 481 AGATGTTGGCGAGTGCACCGGGGCCCGAGGCTATGCTCCACCGCAAGTCCCTGAACA 540  
QY 541 TCTGGGCGCTGATTTCTGGGCACTCCTCATGACCATTTGATTCATCTGTTACTGTTTCG 600  
DB 541 TCTGGGCGCTGATTTCTGGGCACTCCTCATGACCATTTGATTCATCTGTTACTGTTTCG 600  
QY 601 GCTCTGTACAGTTCACATATTTATGTTACAGATTAATACAGAAACCGGGTTACTAGT 660  
DB 601 GCTCTGTACAGTTCACATATTTATGTTACAGATTAATACAGAAACCGGGTTACTAGT 660  
QY 661 AGCGCCCATAGCTGCAACCTTTGCACTCCACTGTGCAATCTGCGCTTGCACGCTGGG 720  
DB 661 AGCGCCCATAGCTGCAACCTTTGCACTCCACTGTGCAATCTGCGCTTGCACGCTGGG 720  
QY 721 GCTGTTCCCGCTGCCCGCTTGGTCTGCTGCCCTTAGATACAGCAGTTTATACCCACACCT 780  
DB 721 GCTGTTCCCGCTGCCCGCTTGGTCTGCTGCCCTTAGATACAGCAGTTTATACCCACACCT 780  
QY 781 GTCTACACTGACATTCATTAAGTG-ACGTGCTTGTGAAAAAACAATAAACCCTGA 839  
DB 781 GTCTACAGTTCATTCATTAAGTGCACTGTGTTGTGAAAAAACAATAAAGGAGGGG 840  
QY 840 GG 841  
DB 841 GG 842

RESULT 3  
ADE25685  
ID ADE25685 standard; cDNA; 1617 BP.  
XX  
AC ADE25685;  
XX  
DT 29-JAN-2004 (first entry)  
XX  
DE Human cDNA differentially expressed in foam cells #89.  
XX  
KW Human; ss; differential expression; foam cell; LPS; lipopolysaccharide;  
KW cardiovascular disease; atherosclerosis.  
XX  
OS Homo sapiens.  
XX  
FN US2003194721-A1.

XX 16-OCT-2003.  
PD 18-SEP-2002; 2002US-00247671.  
XX PF 19-SEP-2001; 2001US-0323784P.  
XX PR (INCY-) INCYTE GENOMICS INC.  
XX PA Mikita T, Shiffman D, Porter JG, Kaser MR;  
PI WPI; 2003-875398/81.  
XX  
XX Combination containing several polynucleotide that are differentially  
XX expressed in foam cells and complements of the polynucleotides, useful  
XX for diagnosing cardiovascular disease or atherosclerosis.  
PS Claim 1; SEQ ID NO 89; 37pp; English.  
XX  
XX The invention relates to a combination comprising several polynucleotides  
XX having any one of 127 sequences (S1) such as the sequence of human  
XX calmodulin gene, human mRNA for KIAA0930 protein, leukotriene A4  
XX hydrolase, human CGI-142 protein mRNA, human K+ channel beat 2 subunit  
XX mRNA, etc., and their complements. The cDNAs are differentially expressed  
XX in LPS (lipopolysaccharide)-treated foam cells. Also included are  
XX obtaining an extended or full length gene from a library of nucleic acid  
XX sequences, an expression vector containing the nucleic acids, a host cell  
XX containing the vector, a purified polypeptide appearing as ADE25750 and  
XX ADE25751, producing a protein by culturing the host cell, and a  
XX composition comprising a purified antibody that specifically binds to the  
XX proteins. The foam cell-expressed nucleic acids are useful for a high  
XX throughput detection of differential expression of one or more  
XX polynucleotides in a sample. The sample is from a subject with  
XX atherosclerosis and comparison with a standard defines early, mid or late  
XX stages of the disorder. The foam cell-expressed nucleic acids are useful  
XX for high throughput screening of a library of molecules or compounds to  
XX identify a ligand which binds a polynucleotide. The library is chosen  
XX from DNA molecules, peptides, proteins and RNA molecules. The protein is  
XX useful for a high throughput screening of library of molecules or  
XX compounds to identify at least one ligand which specifically binds a  
XX protein, for purifying a ligand from a sample for making an antibody. The  
XX foam cell-expressed nucleic acids are useful for diagnosing  
XX cardiovascular disorder. The foam cell-expressed nucleic acids are useful  
XX as elements on a microarray which can be used for detecting related  
XX polynucleotide in a sample, diagnosing cardiovascular disease,  
XX atherosclerosis. The present sequence represents a cDNA whose expression  
XX is upregulated in LPS treated foam cells.  
SQ Sequence 1617 BP; 348 A; 480 C; 445 G; 344 T; 0 U; 0 Other;  
Query Match 63.4%; Score 703; DB 9; Length 1617;  
Best Local Similarity 97.2%; Pred. No. 7.3e-182;  
Matches 726; Conservative 0; Mismatches 20; Indels 1; Gaps 1;  
QY 106 AGCAGGAAATAGAAACTTAAGAGAAATACACACTTCTGAGAACTGAAACGACAGGGGAA 165  
DB 379 AGCAGGAAATAGAAACTTAAGAGAAATACACACTTCTGAGAACTGAAACGACAGGGGAA 438  
QY 166 AGGAGGTCTCACTGAGCACCCTCCAGCATCCGGACACACAGCGGCGCTTCGCTCCACG 225  
DB 439 AGGAGGTCTCACTGAGCACCCTCCAGCATCCGGACACACAGCGGCGCTTCGCTCCACG 498  
QY 226 CAGAAAACACACTTCTCAAACTTCACTCAACACTTCTTCCCAAGCCAGAGATGC 285  
DB 499 CAGAAAACACACTTCTCAAACTTCACTCAACACTTCTTCCCAAGCCAGAGATGC 558  
QY 286 ACAGGAGGACATGAGTGGTGTCTGGGGGACCCCGACACCATCTTCCCAAGGT 345  
DB 559 ACAAGGAGGACATGAGTGGTGTCTGGGGGACCCCGACACCATCTTCCCAAGGT 618  
QY 346 CCACCGTGATCAACATCCACAGCGAGACCTCCGCTGCCCGACCATGTGCTGTGCTCCCTGT 405  
DB 619 CCACCGTGATCAACATCCACAGCGAGACCTCCGCTGCCCGACCATGTGCTGTGCTCCCTGT 678

QY 406 TCACACCCCTCTTCTGAACTGGTCTGTCTGGGTTCTATAGCATTCGGCTACTCCGTGA 465  
 Db 679 TCACACCCCTCTTCTGAACTGGTCTGTCTGGGTTCTATAGCATTCGGCTACTCCGTGA 738  
 QY 466 AGTCTAGGACAGGAGATGTTGGCGAGCTGACCGGGGCCAGGCTATGCTCCACCG 525  
 Db 739 AGTCTAGGACAGGAGATGTTGGCGAGCTGACCGGGGCCAGGCTATGCTCCACCG 798  
 QY 526 CCAAGTGGCTGAACATCTGGGCCCTGATCTGGGCATCTCATGACCATGGATTATCC 585  
 Db 799 CCAAGTGGCTGAACATCTGGGCCCTGATCTGGGCATCTCATGACCATGGATTATCC 858  
 QY 586 TGTTACTGGTATTCGGCTGTGACAGCTACCATATTATGTACAGATAATACAGGAA 645  
 Db 859 TGTTACTGGTATTCGGCTGTGACAGCTACCATATTATGTACAGATAATACAGGAA 918  
 QY 646 AACGGGGTTACTAGTAGCGGCCCATAGCCTGCAACCTTTGCACTCCACTGTGCAATGCTG 705  
 Db 919 AACGGGGTTACTAGTAGCGGCCCATAGCCTGCAACCTTTGCACTCCACTGTGCAATGCTG 978  
 QY 706 GCGCTGACGCTGGGCTGTGCGCCCTGCGCCCTGGTCTGCTGCGCCCTAGATACAGAGTT 765  
 Db 979 GCGCTGACGCTGGGCTGTGCGCCCTGCGCCCTGGTCTGCTGCGCCCTAGATACAGAGTT 1038  
 QY 766 TATACCCACACACCTGTCTACACTGACATTCATATAAAGTG-ACGTGCTTGTGAAAAAAA 824  
 Db 1039 TATACCCACACACCTGTCTACAGTGTCAATCAATAAGTGACAGTGTCTTGTGATGCT 1098  
 QY 825 ACAATAAACCAGGGGGGGGGCGG 851  
 Db 1099 GTGACTTCATGTGAGTGGGGGGGGCTG 1125

RESULT 4  
 AAS95008  
 ID AAS95008 standard; DNA; 1095 BP.  
 AC AAS95008;  
 XX  
 XX  
 DT 14-FEB-2002 (first entry)  
 DE Human DNA sequence #263 expressed during foam cell differentiation.  
 XX  
 XX  
 KW Human; foam cell differentiation; atherosclerosis; cerebral stroke;  
 XX cardiovascular disorder; coronary artery disease; gene therapy; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 FN WO200177389-A2.  
 XX  
 PD 18-OCT-2001.  
 XX  
 PF 04-APR-2001; 2001WO-US011128.  
 XX  
 PR 05-APR-2000; 2000US-0195106P.  
 XX  
 XX (INCY-) INCYTE GENOMICS INC.  
 XX  
 XX Shiffman D, Somogyi R, Lawn R, Seilhamer JJ, Porter GJ, Mikita T;  
 PI Tai J;  
 XX  
 XX WPI; 2002-010925/01.  
 DR  
 XX  
 XX Composition useful for diagnosis of conditions, disorders or diseases  
 PT associated with atherosclerosis, comprises several polynucleotides that  
 PT are differentially expressed in foam cell development.  
 XX  
 XX Claim 1; Page 303; 315pp; English.  
 PS  
 XX The present invention relates to the isolation of human polynucleotide  
 CC sequences that are differentially expressed during foam cell  
 CC differentiation. The polynucleotide sequences of the invention or a

CC composition comprising these polynucleotides are useful as a high  
 CC throughput method for detecting altered expression of one or more  
 CC polynucleotides in a sample. The polynucleotides can be used in the  
 CC diagnosis of disorders associated with foam cell development such as  
 CC atherosclerosis, cerebral stroke, and cardiovascular disorders such as  
 CC coronary artery disease. The polynucleotide sequences can also be used as  
 CC PCR primers and probes. The polynucleotides of the invention are also  
 CC useful in gene therapy. AAS94746-AAS95021 represent the human  
 CC polynucleotide sequences of the invention which are differentially  
 CC expressed during foam cell differentiation  
 XX  
 XX

SQ Sequence 1095 BP; 267 A; 326 C; 268 G; 234 T; 0 U; 0 Other;  
 Query Match 63.1%; Score 699.8; DB 6; Length 1095;  
 Best Local Similarity 99.6%; Pred. No. 4.6e-181;  
 Matches 712; Conservative 0; Mismatches 2; Indels 1; Gaps 1;  
 QY 106 AGCAGGAAATAGAACTTAAGAGAAATACACACTTCTGAGAACTGAAACGACAGGGGAA 165  
 Db 381 AGCAGGAAATAGAACTTAAGAGAAATACACACTTCTGAGAACTGAAACGACAGGGGAA 440  
 QY 166 AGGAGGTCTCACTGAGCACCCTCCAGCATCCGACACACACAGCGGCCCTTCGCTCCACG 225  
 Db 441 AGGAGGTCTCACTGAGCACCCTCCAGCATCCGACACACACAGCGGCCCTTCGCTCCACG 500  
 QY 226 CAGAAACACCACTTCTCAAACTTCACTCAAACTTCTTCCCAAGCGAGAGATGC 285  
 Db 501 CAGAAACACCACTTCTCAAACTTCACTCAAACTTCTTCCCAAGCGAGAGATGC 560  
 QY 286 ACAAGGAGAAATAGAGTGGTGTCTGGGGACACCCCGACACCATCTTCCCAAGGT 345  
 Db 561 ACAAGGAGAAATAGAGTGGTGTCTGGGGACACCCCGACACCATCTTCCCAAGGT 620  
 QY 346 CCACCGTATCAACATCCAACGAGACCTCGGTGCGGACCATGTGCTGTGCTCCCTGT 405  
 Db 621 CCACCGTATCAACATCCAACGAGACCTCGGTGCGGACCATGTGCTGTGCTCCCTGT 680  
 QY 406 TCACACCCCTCTTCTGAACTGGTCTGTGGGCTTATAGCATTCGCTACTCCGTGA 465  
 Db 681 TCACACCCCTCTTCTGAACTGGTCTGTGGGCTTATAGCATTCGCTACTCCGTGA 740  
 QY 466 AGTCTAGGACAGGAGATGTTGGCGAGCTGACCGGGGCCAGGCTATGCTCCACCG 525  
 Db 741 AGTCTAGGACAGGAGATGTTGGCGAGCTGACCGGGGCCAGGCTATGCTCCACCG 800  
 QY 526 CCAAGTGGCTGAACATCTGGGCCCTGATTCGCGCATCTCATGACCATTCGATTCATCC 585  
 Db 801 CCAAGTGGCTGAACATCTGGGCCCTGATTCGCGCATCTCATGACCATTCGATTCATCC 860  
 QY 586 TGTTACTGGTATTCGGCTCTGTGACAGTCTACCATATTATGTACAGATAATACAGGAA 645  
 Db 861 TGTTACTGGTATTCGGCTCTGTGACAGTCTACCATATTATGTACAGATAATACAGGAA 920  
 QY 646 AACGGGGTTACTAGTAGCGGCCCATAGCCTGCAACCTTTGCACTCCACTGTGCAATGCTG 705  
 Db 921 AACGGGGTTACTAGTAGCGGCCCATAGCCTGCAACCTTTGCACTCCACTGTGCAATGCTG 980  
 QY 706 GCGCTGACGCTGGGCTGTGCGCCCTGCGCCCTGGTCTGCTGCGCCCTAGATACAGAGTT 765  
 Db 981 GCGCTGACGCTGGGCTGTGCGCCCTGCGCCCTGGTCTGCTGCGCCCTAGATACAGAGTT 1040  
 QY 766 TATACCCACACACCTGTCTACACTGACATTCATATAAAGTG-ACGTGCTTGTGAAA 819  
 Db 1041 TATACCCACACACCTGTCTACAGTGTCAATCAATAAAGTGACAGCTGTGTTGAAA 1095

RESULT 5  
 ABK46192  
 ID ABK46192 standard; cDNA; 842 BP.  
 XX  
 AC ABK46192;  
 XX  
 DT 05-JUN-2002 (first entry)

XX cDNA encoding colon tumour protein, SEQ ID No 1743.  
 XX Human; colon tumour; vaccine; colon cancer; immunogenic; immunotherapy;  
 KW gene; ss.  
 XX Homo sapiens.  
 OS  
 XX WO200212328-A2.  
 PN  
 PD 14-FEB-2002.  
 XX  
 XX 31-JUL-2001; 2001WO-US024218.  
 PF  
 XX 03-AUG-2000; 2000US-0223283P.  
 PR  
 XX 28-MAR-2001; 2001US-0279783P.  
 PR  
 XX 29-JUN-2001; 2001US-0302051P.  
 XX  
 XX (CORI-) CORIXA CORP.  
 PA  
 XX King GE, Meagher MJ, Xu J, Secretist H;  
 PI  
 XX WPI; 2002-241739/29.  
 DR  
 XX New colon cancer polypeptides and polynucleotides, useful as vaccines,  
 PT for diagnosing, preventing, and treating colon cancer, and as markers for  
 PT the progression of cancer.  
 PT  
 XX Claim 1; SEQ ID NO 1743; 147pp; English.  
 PS  
 XX The invention relates to polynucleotides encoding colon tumour proteins.  
 CC The polynucleotides and encoded polypeptides are useful in pharmaceutical  
 CC compositions, such as vaccines, for the diagnosis, prevention, and  
 CC treatment of colon cancer. Polynucleotide sequences may be used as  
 CC hybridisation probes or primers, and in the design and preparation of  
 CC ribozyme molecules for inhibiting expression of tumour polypeptides and  
 CC proteins in tumour cells. The compositions are useful for stimulating an  
 CC immune response against cancer, particularly for the immunotherapy of  
 CC colon cancer, and as markers for the progression of cancer. ABK4450-  
 CC ABK46237 represent coding sequences of human colon tumour proteins of the  
 CC invention. Note: With the exception of SEQ ID No 1 and 2, the sequence  
 CC data for this patent did not form part of the printed specification but  
 CC was supplied by the European Patent Office  
 XX  
 XX Sequence 842 BP; 206 A; 264 C; 180 G; 192 T; 0 U; 0 Other;  
 SQ  
 Query Match 62.8%; Score 696.2; DB 6; Length 842;  
 Best Local Similarity 99.4%; Pred. No. 4e-180;  
 Matches 709; Conservative 0; Mismatches 3; Indels 1; Gaps 1;  
 106 AGCAGGAATAGAACTTAAGAGAAATACACACTTCTGAGAACTGAAACGACAGGGGAA 165  
 130 AGCAGGAATAGAACTTAAGAGAAATACACACTTCTGAGAACTGAAACGACAGGGGAA 189  
 166 AGGAGGTCTCACTGAGCAACCGTCCAGCATCCGGACACACACAGCGCGCTTGGCTCCAG 225  
 190 AGGAGGTCTCACTGAGCAACCGTCCAGCATCCGGACACACACAGCGCGCTTGGCTCCAG 249  
 226 CAGAAACCACTTCTCAAACTTCACTCAACACTTCTTCCCAAGCCAGCAATGC 285  
 250 CAGAAACCACTTCTCAAACTTCACTCAACACTTCTTCCCAAGCCAGCAATGC 309  
 286 ACAGGAGGAACATGAGTGGTGTGCTGGGGGACCCCGCCAGCAATCCTTCCAAAGGT 345  
 310 ACAGGAGGAACATGAGTGGTGTGCTGGGGGACCCCGCCAGCAATCCTTCCAAAGGT 369  
 346 CCACCGTATCAACTCCACAGCGAGACTCCGTCGCGCCGACCATGCTGCTGCTCCCTGT 405  
 370 CCACCGTATCAACTCCACAGCGAGACTCCGTCGCGCCGACCATGCTGCTGCTCCCTGT 429  
 406 TCAACACCTCTCTTGAACCTGGTGTGCTGCTGGGCTTCATAGCATTCGGCTACTCCGTGA 465  
 430 TCAACACCTCTCTTGAACCTGGTGTGCTGCTGGGCTTCATAGCATTCGGCTACTCCGTGA 489

QY 466 AGTCTAGGACAGGAAGATGTTGGGACCTGACGGGGCCCGAGCCCTATGCTCCACCG 525  
 DB 490 AGTCTAGGACAGGAAGATGTTGGGACCTGACGGGGCCCGAGCCCTATGCTCCACCG 549  
 QY 526 CCAAGTGCCTGAACATCTGGGCGCTGATTTCTGGGCATCCTCATGACCAATGGATTCATCC 585  
 DB 550 CCAAGTGCCTGAACATCTGGGCGCTGATTTCTGGGCATCCTCATGACCAATGGATTCATCC 609  
 QY 586 TGTACTGTGTTATTCGGCTCTGTGACAGCTTACCATATTTATTTACAGATAATACAGGAA 645  
 DB 610 TGTACTGTGTTATTCGGCTCTGTGACAGCTTACCATATTTATTTACAGATAATACAGGAA 669  
 QY 646 AACGGGTTACTAGTAGTACCGCCCATAGCTCAACCTTTGCACTCCACTGTGCAATGCTG 705  
 DB 670 AACGGGTTACTAGTAGTACCGCCCATAGCTCAACCTTTGCACTCCACTGTGCAATGCTG 729  
 QY 706 GCCTGTGACCGTGGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 765  
 DB 730 GCCTGTGACCGTGGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 789  
 QY 766 TATACCCACACACCTGTCTACACTGACATTCATAAAGTG-ACGTGCTTGTGA 817  
 DB 790 TATACCCACACACCTGTCTACACTGACATTCATAAAGTGACGCTGCTTGTGA 842  
 RESULT 6  
 ABA83110  
 ID ABA83110 standard; DNA; 853 BP.  
 XX  
 AC ABA83110;  
 XX  
 DT 08-FEB-2002 (first entry)  
 XX  
 DE IFN-induced transmembrane protein 1 ovarian tumour marker gene, #57.  
 XX  
 KW Ovarian tumour marker gene; human; overexpression; upregulation;  
 KW epithelial tumour; cancer; diagnosis; prognosis; disease monitoring;  
 KW identification; serous cystadenoma; borderline serous tumour;  
 KW serous cystadenocarcinoma; mucinous cystadenocarcinoma;  
 KW mucinous cystadenoma; borderline mucinous tumour; endometrioid carcinoma;  
 KW undifferentiated carcinoma; clear cell adenocarcinoma; cystadenofibroma;  
 KW adenofibroma; Brenner tumour; serial analysis of gene expression; SAGE;  
 KW immune response pathway; cell proliferation regulation; protein folding;  
 KW membrane localised; secreted; therapeutic target; cytostatic;  
 KW gene therapy; vaccine; ds.  
 XX  
 XX Homo sapiens.  
 OS  
 XX WO200175177-A2.  
 PN  
 XX 11-OCT-2001.  
 PD  
 XX 03-APR-2001; 2001WO-US010947.  
 PF  
 XX 03-APR-2000; 2000US-0194336P.  
 PR  
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.  
 PA  
 XX Morin PJ, Sherman-Baust CA, Pizer ES, Hough CD;  
 PI  
 XX WPI; 2001-626450/72.  
 DR  
 XX P-PSDB; ABB50284.  
 DR  
 XX Detecting and identifying ovarian tumor, identifying increased risk for  
 PT developing ovarian cancer, and determining effectiveness of ovarian  
 PT cancer treatment, by measuring expression level of ovarian tumor marker  
 PT gene.  
 XX  
 PS Claim 23; Page 102; 140pp; English.  
 XX  
 XX The invention relates to methods for diagnosing and prognosing ovarian  
 CC tumours in an individual via the detection and measurement of the

CC	expression of ovarian tumour marker genes (ABA83081-ABA83122, ABA83180, ABA83182 and ABA83184) or segments thereof (ABA83123-ABA83169, ABA83179, ABA83181 and ABA83183). The methods of the invention are useful for detecting an ovarian tumour in a patient, for identifying an individual at increased risk for developing ovarian cancer, in prognostic tests for assessing the relative severity of ovarian cancer, in tests for monitoring a patient in remission from ovarian cancer and in tests for monitoring disease status in a patient being treated for ovarian cancer. The methods can additionally be used to identify a particular tumour as being an ovarian tumour (i.e., an epithelial ovarian tumour selected from serous cystadenoma, borderline serous tumour, serous cystadenocarcinoma, mucinous cystadenoma, borderline mucinous tumour, mucinous cystadenocarcinoma, endometrioid carcinoma, undifferentiated carcinoma, clear cell adenocarcinoma, cystadenofibroma, adenofibroma and Brenner tumour). The ovarian tumour marker genes of the invention were identified using SAGE (serial analysis of gene expression) and were found to be overexpressed in a broad variety of ovarian epithelial tumour cells relative to normal ovarian epithelial cells. The marker genes are implicated in immune response pathways, and many of these are membrane-proliferation and in protein folding, and many of these are membrane-localised or secreted. In addition to their use as diagnostic and prognostic markers, the ovarian tumour marker genes or their encoded proteins may be used as therapeutic targets for the treatment and prevention of ovarian cancer. Sequences ABA83081-ABA83122, ABA83180, ABA83182 and ABA83184 represent the ovarian tumour marker genes of the invention	CC
CC	Sequence 853 BP; 208 A; 268 C; 182 G; 195 T; 0 U; 0 Other;	XX
CC	Query Match 62.8%; Score 696.2; DB 5; Length 853;	XX
CC	Best Local Similarity 99.4%; Pred. No. 4e-180;	XX
CC	Matches 709; Conservative 0; Mismatches 3; Indels 1; Gaps 1;	XX
QY	106 AGCAGGAAATAGAAACTTAAGAGAAATACACACTTCTGAGAAACTGAAACGACAGGGGAA 165	QY
DB	141 AGCAGGAAATAGAAACTTAAGAGAAATACACACTTCTGAGAAACTGAAACGACAGGGGAA 200	DB
QY	166 AGGAGTCTCACTGAGCACCGTCCGAGCATCGGACACACAGCGGGCCCTTCGTCTCCACG 225	QY
DB	201 AGGAGGTCTCACTGAGCACCGTCCGAGCATCGGACACACAGCGGGCCCTTCGTCTCCACG 260	DB
QY	226 CAGAAAAACCACTTCTCAAAACCTTCACTCAACACTTCTTCCCCAAAGCCAGAAGATGC 285	QY
DB	261 CAGAAAAACCACTTCTCAAAACCTTCACTCAACACTTCTTCCCCAAAGCCAGAAGATGC 320	DB
QY	286 ACAGGAGGACATGAGTGCTGTGCTGGGGGACCCGCCAGACCACTCTTCCAAAGT 345	QY
DB	321 ACAGGAGGAAACATGAGTGCTGTGCTGGGGGACCCGCCAGACCACTCTTCCAAAGT 380	DB
QY	346 CCACCGTGATCAACATCCACAGCGAGACCTCGTGCCCGCAACATCGTCTGGTCCCTGT 405	QY
DB	381 CCACCGTGATCAACATCCACAGCGAGACCTCGTGCCCGCAACATCGTCTGGTCCCTGT 440	DB
QY	406 TCAACACCCCTCTTCTGAACTGGTCTGTGGGCTTTCATAGCATTCGCCTACTCCGTGA 465	QY
DB	441 TCAACACCCCTCTTCTGAACTGGTCTGTGGGCTTTCATAGCATTCGCCTACTCCGTGA 500	DB
QY	466 AGTCTAGGGACAGGAAGATGTTGGCGAGCTGACCGGGGCCAGGCGCTATGCCTCCACCG 525	QY
DB	501 AGTCTAGGGACAGGAAGATGTTGGCGAGCTGACCGGGGCCAGGCGCTATGCCTCCACCG 560	DB
QY	526 CCAAGTGCTGAACTCTGGGCCCTGATTCCTGGGCATCTCATGACCATTTGATTCATCC 585	QY
DB	561 CCAAGTGCTGAACTCTGGGCCCTGATTCCTGGGCATCTCATGACCATTTGATTCATCC 620	DB
QY	586 TGTTACTGGTATTCGGCTCTGTGACAGTCTACCATATATTGTTACAGATAATACAGGAAA 645	QY
DB	621 TGTCAC TGTTACTGGTCTGTGACAGTCTACCATATATTGTTACAGATAATACAGGAAA 680	DB
QY	646 AACGGGGTTACTAGTAGCCGCCCATAGCTGCAACCTTTGCACTCCACTGTGCAATGCTG 705	QY
DB	681 AACGGGGTTACTAGTAGCCGCCCATAGCTGCAACCTTTGCACTCCACTGTGCAATGCTG 740	DB

XX exposure of a subject to a pathogen or sterile inflammatory disease (e.g. psoriasis, rheumatoid arthritis, glomerulonephritis, asthma, thrombosis, cardiac reperfusion injury, renal reperfusion injury, ARDS, adult respiratory distress syndrome, inflammatory bowel disease, Crohn's disease, ulcerative colitis, periodontal disease; also bacterial infection, viral infection, parasitic infection, protozoal infection, fungal infection and MS is useful for treating one of the above conditions. The present sequence represents a gene differentially expressed in granulocytes. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 853 BP; 208 A; 268 C; 182 G; 195 T; 0 U; 0 Other;  
Query Match 62.8%; Score 696.2; DB 6; Length 853;  
Best Local Similarity 99.4%; Pred. No. 4e-180;  
Matches 709; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

2Y 106 AGCAGGAATAGAACTTAAGAGAAATACACACTTCTGAGAACTGAAACGACAGGGGAA 165  
Db 141 AGCAGGAATAGAACTTAAGAGAAATACACACTTCTGAGAACTGAAACGACAGGGGAA 200  
Qy 166 AGGAGGTCTCACTGAGCAGCGTCCAGCATCCGAGACACACAGCGGCCCTTCGCTCCAG 225  
Db 201 AGGAGGTCTCACTGAGCAGCGTCCAGCATCCGAGACACACAGCGGCCCTTCGCTCCAG 260  
Qy 226 CAGAAACACACTTCTCAAACTTCACTCAACACTTCTCCCAAGCCAGAGATGC 285  
Db 261 CAGAAACACACTTCTCAAACTTCACTCAACACTTCTCCCAAGCCAGAGATGC 320  
Qy 286 ACAAGGAGGAACATGAGGTGGTGTCTGGGGACACCCCGACCATCTTCCAAAGGT 345  
Db 321 ACAAGGAGGAACATGAGGTGGTGTCTGGGGACACCCCGACCATCTTCCAAAGGT 380  
Qy 346 CCACCGTGATCAATCCACAGCAGACTCCGTGCGCGACCATGTCTGTGTCCTGT 405  
Db 381 CCACCGTGATCAATCCACAGCAGACTCCGTGCGCGACCATGTCTGTGTCCTGT 440  
Qy 406 TCAACACCTCTCTTGAACCTGTGTCTGTGGGCTTTCATAGCAATCGCTTCTCCGTGA 465  
Db 441 TCAACACCTCTCTTGAACCTGTGTGTCTGTGGGCTTTCATAGCAATCGCTTCTCCGTGA 500  
Qy 466 AGTCTAGGACAGGAATGTTGGCAGCTGACCGGGGCCAGGCCCTATGCTCCACCG 525  
Db 501 AGTCTAGGACAGGAATGTTGGCAGCTGACCGGGGCCAGGCCCTATGCTCCACCG 560  
Qy 526 CCAAGTGCTTGAACCTGTGGGCTGTCTGGGCATCTTCATGACCATTCGATTCATCC 585  
Db 561 CCAAGTGCTTGAACCTGTGGGCTGTCTGGGCATCTTCATGACCATTCGATTCATCC 620  
Qy 586 TGTACTGTGTTTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGG 645  
Db 621 TGTACTGTGTTTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGG 680  
Qy 705 AACGGGGTACTAGTACCGCCCTAGCTGCAACCTTTGCACTCCACTGTGCAATGCTG 705  
Db 681 AACGGGGTACTAGTACCGCCCTAGCTGCAACCTTTGCACTCCACTGTGCAATGCTG 740  
Qy 706 GCCTGTGACGCTGGGGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGG 765  
Db 741 GCCTGTGACGCTGGGGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGG 800  
Qy 766 TATACCCACACACTGTCTACACTGACATTCATTAATAAGTG-ACGTCCTGTGCA 817  
Db 801 TATACCCACACACTGTCTACAGTGTCTATTAATAAGTGACGTCCTGTGCA 853

RESULT 8  
ABK64497  
ID ABK64497 standard; DNA; 853 BP.  
XX  
AC ABK64497;

XX 18-JUN-2002 (first entry)  
Dt Human benign prostatic hyperplasia gene #392.  
De Human; benign prostatic hyperplasia; BPH; prostate cancer; gene; ds.  
Xx Homo sapiens.  
Xx WO200212440-A2.  
Xx 14-FEB-2002.  
Xx 07-AUG-2001; 2001WO-US024708.  
Xx 07-AUG-2000; 2000US-0223323P.  
Xx 05-JUN-2001; 2001US-00873319.  
Xx (GENE-) GENE LOGIC INC.  
Xx (NLSB ) JAPAN TOBACCO INC.  
Xx Munger WE, Kulkarni P, Getzenberg RH, Waga I, Yamamoto J;  
Xx WPI; 2002-257476/30.  
Xx Identifying drugs for and diagnosing benign prostatic hyperplasia, by detecting expression levels of one or more genes in prostate cells from patient that are differentially regulated compared to normal prostate cells.  
Xx Disclosure; Page 222; 444pp; English.  
Xx The invention relates to a method of diagnosing (I) the onset or progression of benign prostatic hyperplasia (BPH), or screening (II) for or identifying an agent that modulates the onset or progression of BPH. The method is based on changes in gene expression in BPH tissue isolated from patients exhibiting different clinical states of prostate hyperplasia as compared to normal prostate tissue. (I) comprises detecting the expression levels of one or more genes in prostate cells from the subject that are differentially regulated compared to normal prostate cells. (II) comprises preparing a first gene expression profile of BPH cells or BPH-like cell population, exposing the cells to the agent, and preparing a second gene expression profile of the agent exposed cells, and comparing the first and second gene expression profiles. (I) is useful for diagnosing the onset or progression of BPH. (II) is useful for identifying an agent that modulates the onset or progression of BPH. The methods are useful to present information identifying the expression level in a tissue or cells, by comparing the expression level of genes given in the specification in the tissue or cells to the level of expression of gene in the database, and displaying the expression levels of at least one gene in the tissue or cell sample compared to the expression level in BPH. Agents using (II) are useful for treating BPH or prostate cancer. ABK64106-ABK64860 represent human benign prostatic hyperplasia gene sequences of the invention

XX Sequence 853 BP; 208 A; 268 C; 182 G; 195 T; 0 U; 0 Other;  
Query Match 62.8%; Score 696.2; DB 6; Length 853;  
Best Local Similarity 99.4%; Pred. No. 4e-180;  
Matches 709; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

Qy 106 AGCAGGAATAGAACTTAAGAGAAATACACACTTCTGAGAACTGAAACGACAGGGGAA 165  
Db 141 AGCAGGAATAGAACTTAAGAGAAATACACACTTCTGAGAACTGAAACGACAGGGGAA 200  
Qy 166 AGGAGGTCTCACTGAGCAGCGTCCAGCATCCGAGACACACAGCGGCCCTTCGCTCCAG 225  
Db 201 AGGAGGTCTCACTGAGCAGCGTCCAGCATCCGAGACACACAGCGGCCCTTCGCTCCAG 260  
Qy 226 CAGAAACACACTTCTCAAACTTCACTCAACACTTCTCCCAAGCCAGAGATGC 285  
Db 261 CAGAAACACACTTCTCAAACTTCACTCAACACTTCTCCCAAGCCAGAGATGC 320

286 ACAAGAGGAAACATGAGTGGCTGTGCTGGGGGACACCCAGCAGCACCATCTTCCAAAGT 345  
321 ACAGAGGAGACATGAGTGGCTGTGCTGGGGGACACCCAGCAGCACCATCTTCCAAAGT 380  
346 CAACCGTGATCAACATCCACAGGAGACCTCCGTCGGGACCATGTGCTGGTCCCTGT 405  
381 CCACCGTGATCAACATCCACAGGAGACCTCCGTCGGGACCATGTGCTGGTCCCTGT 440  
406 TCAACACCCCTCTTCTTGAACCTGGTGTCTGGGCTTCATAGCATTCGCCCTACTCCGTGA 465  
441 TCAACACCCCTCTTCTTGAACCTGGTGTCTGGGCTTCATAGCATTCGCCCTACTCCGTGA 500  
466 AGTCTAGGAGACAGAAAGATGGTTGGGACGTGACCGGGGCCAGCCCTATGCTCCACCG 525  
501 AGTCTAGGAGACAGAAAGATGGTTGGGACGTGACCGGGGCCAGCCCTATGCTCCACCG 560  
526 CCAAGTGCCCTGAACATCTGGGACCTGATCTGGGACCTCTCATAGCATTCGCCCTACTCCGTGA 585  
561 CCAAGTGCCCTGAACATCTGGGACCTGATCTGGGACCTCTCATAGCATTCGCCCTACTCCGTGA 620  
586 TGTACTGTGATTTGGCTCTGTGACAGTCTACCATATTTATGTTACAGATAATACAGGAAA 645  
621 TGTACTGTGATTTGGCTCTGTGACAGTCTACCATATTTATGTTACAGATAATACAGGAAA 680  
646 AACGGGGTTACTAGTAGCGGCCCATAGCTGCAACCTTTGCACTCCACTGTGCAATGCTG 705  
681 AACGGGGTTACTAGTAGCGGCCCATAGCTGCAACCTTTGCACTCCACTGTGCAATGCTG 740  
706 GCCTTGACACGCTGGGGCTGTTGCCCTCCCTGGTCTGGTCCCTAGATACAGAGTT 765  
741 GCCTTGACACGCTGGGGCTGTTGCCCTCCCTGGTCTGGTCCCTAGATACAGAGTT 800  
766 TATACCCACACACCTGTCTACAGTGTCTCAATTAAGTGACGCTGTGTGA 817  
801 TATACCCACACACCTGTCTACAGTGTCTCAATTAAGTGACGCTGTGTGA 853

RESULT 9

ABN95610  
ID ABN95610 standard; DNA; 853 BP.

AC AC  
XX XX  
ABN95610;

13-AUG-2002 (first entry)

Gene #2108 used to diagnose liver cancer.

Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;  
metastatic liver tumor; cytostatic; expression profile; disease state;  
disease progression; drug toxicity; drug efficacy; drug metabolism.

Homo sapiens.

W0200229103-A2.

11-APR-2002.

02-OCT-2001; 2001WO-US030589.

02-OCT-2000; 2000US-0237054P.

(GENE-) GENE LOGIC INC.

Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;

WPI; 2002-426119/45.

Diagnosing and detecting the progression of liver cancer, hepatocellular  
carcinoma or metastatic liver tumor in a patient, involves detecting the  
level of expression of two or more genes in a liver tissue sample.

Claim 1; SEQ ID NO 2108; 298pp; English.

RESULT 10  
ABT31907

CC The invention relates to a novel method for diagnosing and detecting the  
CC progression of liver cancer, hepatocellular carcinoma or metastatic liver  
CC tumor in a patient, and differentiating metastatic liver cancer from  
CC hepatocellular carcinoma in a patient, involving detecting the level of  
CC expression of two or more genes represented in ABN93503-ABN97455 in a  
CC tissue sample. The method of the invention has hepatotropic, and  
CC cytostatic activity. The method is useful for diagnosing and detecting  
CC the progression of liver cancer, hepatocellular carcinoma and metastatic  
CC liver carcinoma in a patient. The method is useful for identifying  
CC expression profiles which serve as useful diagnostic markers as well as  
CC markers that can be used to monitor disease states, disease progression,  
CC drug toxicity, drug efficacy and drug metabolism. Note: The sequence data  
CC for this patent did not form part of the printed specification, but was  
CC obtained in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences

XX  
SQ Sequence 853 BP; 208 A; 268 C; 182 G; 195 T; 0 U; 0 Other;

Query Match 62.8%; Score 696.2; DB 6; Length 853;

Best Local Similarity 99.4%; Pred. No. 4e-180;

Matches 709; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 106 AGCAGGAATAGAACTTAGAGAAATACACACTTCTGAGAACTGAACGACAGGGGAA 165  
DB 141 AGCAGGAATAGAACTTAGAGAAATACACACTTCTGAGAACTGAACGACAGGGGAA 200  
QY 166 AGGAGGTCTCACTGAGCACCCTCCAGCATCCGACACACAGCGGCCCTTCGCTCCAG 225  
DB 201 AGGAGGTCTCACTGAGCACCCTCCAGCATCCGACACACAGCGGCCCTTCGCTCCAG 260  
QY 226 CAGAAACCACTTCTCAACCTTCACTCAACCTTCTTCCCAAGCCAGAAATGC 285  
DB 261 CAGAAACCACTTCTCAACCTTCACTCAACCTTCTTCCCAAGCCAGAAATGC 320  
QY 286 ACAAGGAGGAACATGAGTGGCTGTGCTGGGGACACCCCGACACCATCTTCCAAAGT 345  
DB 321 ACAAGGAGGAACATGAGTGGCTGTGCTGGGGACACCCCGACACCATCTTCCAAAGT 380  
QY 346 CCACCGTGATCAACATCCACAGCGAGACCTCGTGGCCCGACCATGCTGCTGCTCCTGT 405  
DB 381 CCACCGTGATCAACATCCACAGCGAGACCTCGTGGCCCGACCATGCTGCTGCTCCTGT 440  
QY 406 TCAACACCCCTCTTCTTGAACCTGGTGTCTGTGGGCTTCATAGCATTCGCCCTACTCCGTGA 465  
DB 441 TCAACACCCCTCTTCTTGAACCTGGTGTCTGTGGGCTTCATAGCATTCGCCCTACTCCGTGA 500  
QY 466 AGTCTAGGAGACAGAAAGATGGTTGGGACGTGACCGGGGCCAGGCTATGCCCTCCACCG 525  
DB 501 AGTCTAGGAGACAGAAAGATGGTTGGGACGTGACCGGGGCCAGGCTATGCCCTCCACCG 560  
QY 526 CCAAGTGCCCTGAACATCTGGGACCTGATCTGGGACCTCTCATAGCATTCGCCCTACTCC 585  
DB 561 CCAAGTGCCCTGAACATCTGGGACCTGATCTGGGACCTCTCATAGCATTCGCCCTACTCC 620  
QY 586 TGTACTGTGATTTGGCTCTGTGACAGTCTACCATATTTATGTTACAGATAATACAGGAAA 645  
DB 621 TGTACTGTGATTTGGCTCTGTGACAGTCTACCATATTTATGTTACAGATAATACAGGAAA 680  
QY 646 AACGGGGTTACTAGTAGCGGCCCATAGCTGCAACCTTTGCACTCCACTGTGCAATGCTG 705  
DB 681 AACGGGGTTACTAGTAGCGGCCCATAGCTGCAACCTTTGCACTCCACTGTGCAATGCTG 740  
QY 706 GCCTTGACACGCTGGGGCTGTTGCCCTCCCTGGTCTGGTCCCTAGATACAGAGTT 765  
DB 741 GCCTTGACACGCTGGGGCTGTTGCCCTCCCTGGTCTGGTCCCTAGATACAGAGTT 800

QY 766 TATACCCACACACCTGTCTACAGTGTCTCAATTAAGTGACGCTGTGTGA 817

DB 801 TATACCCACACACCTGTCTACAGTGTCTCAATTAAGTGACGCTGTGTGA 853











CC	proteins. The DNA and protein sequences of the invention are useful as:		
CC	markers for tissues in which the corresponding protein is preferentially		
CC	expressed; as molecular weight markers on gels; as chromosome markers or		
CC	tags; to identify chromosomes or to map related gene positions; and to		
CC	compare with endogenous DNA sequences in patients to identify potential		
CC	genetic disorders. The present DNA sequence represents a gene of the		
CC	invention.		
XX			
SQ	Sequence 651 BP; 153 A; 205 C; 151 G; 142 T; 0 U; 0 Other;		
	Query Match 49.9%; Score 553.4; DB 9; Length 651;		
	Best Local Similarity 94.5%; Pred. No. 4.4e-141;		
	Matches 596; Conservative 0; Mismatches 31; Indels 4; Gaps 2;		
QY	157 ACAGGGGAAGAGGTCTC	ACTGAGCACCCTCCAGCATCCGACACACACAGCGGCCCTT	216
DB	3 ACGTGGGAAGAGGTCT	CAC	62
QY	217 CGCTCCACGACAGAAAC	CCACCTTCAACCTTCACTCAACCTTCCCTCCCAAGCC	276
DB	63 CGCTCCACGACAGAAAC	CCACCTTCAACCTTCACTCAACCTTCCCTCCCAAGCC	122
QY	277 AGAAGATGCACAGAGGAG	GAATGAGTGGCTGTGTGGGGGACACCCCGACACCATCC	336
DB	123 AGAAGATGCACAGAGGAG	GAATGAGTGGCTGTGTGGGGGACACCCCGACACCATCC	182
QY	337 TTCCAAGTCCACCGTGAT	CAACATCCACAGCAGACCTCCGTGCCCGACCATGTGCT	396
DB	183 TTCCAAGTCCACCGTGAT	CAACATCCACAGCAGACCTCCGTGCCCGACCATGTGCT	242
QY	397 GGTCCCTGTTCAACACCC	CTTGAACCTGGTGTGTGTGGGCTTCATAGCATTGCGCT	456
DB	243 GGTCCCTGTTCAACACCC	CTTGAACCTGGTGTGTGTGGGCTTCATAGCATTGCGCT	302
QY	457 ACTCCGTGAAGTCTAGGG	ACAGAGATGGTTGGCGACGTGAACCGGGGCCAGGCCCTATG	516
DB	303 ACTCCGTGAAGTCTAGGG	ACAGAGATGGTTGGCGACGTGAACCGGGGCCAGGCCCTATG	362
QY	517 CCTCCACCGCCCAAGTCC	CTGAACATCTGGGCCCTGATTCTGGGCATCCTCATGACCATTG	576
DB	363 CCTACACCCGACAAAGT	GCCTGAACATCTGGGCCCTGAACTCTGGGCATCTCATGACCATTG	422
QY	577 GATTCACTCTGTTACTGG	TATTCTGGCTCTGTGACAGTCTACCATATTATGTTACAGATAA	636
DB	423 GATTCACTCTGTTACTGG	TATTCTGGCTCTGTGACAGTCTACCATATTATGTTACAGATAA	482
QY	637 TACAGGAAAACCGGGT	TACTAGTACGCCCTAGCTGCAACCTTTCACCTCCACTGT	696
DB	483 TACAGG---AACTGAGT	TACTAGTACGCCCTAGCTGCAACCTTTCACCTCCACTGT	539
QY	697 GCAATGCTGGCCCTGCAC	CGCTGGGGCTGTTGGCCCTGCCCCCTTGGTCTGCCCCCTAGAT	756
DB	540 GCAATGCTTGACCTGCAC	CGCTGAAGCTTGTGCCCTTGCCACTTGGTACTGCGCACTAGAT	599
QY	757 ACAGCAGTTTATACCCAC	CACACCTGTCTACA	787
DB	600 ACAGCAG-TGATACCCAC	CACACCTGTCTTACA	629

Search completed: March 10, 2004, 11:00:17  
Job time : 273.909 secs





Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	351	31.7	352	4	US-09-702-705-1204	Sequence 1204, Ap
2	351	31.7	352	4	US-09-736-457-1204	Sequence 1204, Ap
3	351	31.7	352	4	US-09-614-124B-1204	Sequence 1204, Ap
C 5	351	31.7	352	4	US-09-671-325-1204	Sequence 1204, Ap
	335.8	30.3	764	3	US-08-991-789A-283	Sequence 283, App
C 6	335.8	30.3	764	4	US-09-062-451-283	Sequence 283, App
C 7	335.8	30.3	764	4	US-09-289-198-283	Sequence 283, App
C 8	335.8	30.3	764	4	US-09-429-753-283	Sequence 283, App
9	281.8	25.4	808	4	US-09-023-655-1279	Sequence 1279, Ap
10	175.6	24.8	642	4	US-09-370-838-119	Sequence 119, App
11	175.6	15.8	207	4	US-09-702-705-998	Sequence 998, App
12	175.6	15.8	207	4	US-09-736-457-998	Sequence 998, App
13	175.6	15.8	207	4	US-09-614-124B-998	Sequence 998, App
14	175.6	15.8	207	4	US-09-671-325-998	Sequence 998, App
15	75	6.8	907	4	US-09-620-312D-652	Sequence 652, App
16	48	4.3	1285	4	US-09-149-476-146	Sequence 146, App
17	46.4	4.2	4177	3	US-09-023-082A-23	Sequence 23, App
18	46.4	4.2	4177	4	US-09-348-998-23	Sequence 23, App
19	43	3.9	1434	4	US-09-252-991A-476	Sequence 476, App
C 20	43	3.9	1995	4	US-09-252-991A-398	Sequence 398, App
21	42.6	3.8	1517	4	US-09-304-615-64	Sequence 64, App
22	42.4	3.8	1088	4	US-09-148-545-132	Sequence 132, App
23	42.4	3.8	1143	4	US-09-148-545-79	Sequence 79, App
24	41.6	3.8	1002	4	US-09-457-179-27	Sequence 27, App
25	41.4	3.7	2471	4	US-09-148-545-41	Sequence 41, App
26	41	3.7	433	4	US-09-388-143-34	Sequence 34, App
27	40.8	3.7	319	3	US-09-165-264-8	Sequence 8, App

QY 460 CCGTGAAGTCTAGGACAGGAAGATGGTTGGCGACGTGACCGGGGCCAGG 510  
 Db 301 CCGTGAAGTCTAGGACAGGAAGATGGTTGGCGACGTGACCGGGGCCAGG 351

RESULT 2

US-09-736-457-1204  
 ; Sequence 1204, Application US/09736457  
 ; Patent No. 6509448  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Wang, Tongtong  
 ; APPLICANT: Bangur, Chaitanya S.  
 ; APPLICANT: Lodes, Michael A.  
 ; APPLICANT: Fang, Gary  
 ; APPLICANT: Vedvick, Tom  
 ; APPLICANT: Carter, Darrick  
 ; APPLICANT: Retter, Marc  
 ; APPLICANT: Mannion, Jane  
 ; APPLICANT: Fan, Liqun  
 ; APPLICANT: Wang, AiJun  
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
 ; FILE REFERENCE: 210121.478C15  
 ; CURRENT APPLICATION NUMBER: US/09/736.457  
 ; CURRENT FILING DATE: 2000-12-13  
 ; NUMBER OF SEQ ID NOS: 1864  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 1204  
 ; LENGTH: 352  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapien  
 US-09-736-457-1204

Query Match 31.7%; Score 351; DB 4; Length 352;  
 Best Local Similarity 100.0%; Pred. No. 1.3e-94;  
 Matches 351; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 160 GGGGAAGAGGAGTCTCACTGAGCACCGTCCAGCATCCGACACACAGCGGCCCTTCGC 219  
 Db 1 GGGGAAGAGGAGTCTCACTGAGCACCGTCCAGCATCCGACACACAGCGGCCCTTCGC 60  
 QY 220 TCCACGAGAAAACACACACTTCTCAAACTTCACTCAACACTTCTTCCCAAAGCCAGA 279  
 Db 61 TCCACGAGAAAACACACACTTCTCAAACTTCACTCAACACTTCTTCCCAAAGCCAGA 120  
 QY 280 AGATGCACAGGAGGAACATGAGTGGTGTGCTGGGGGACCCCGACACCATCTTTC 339  
 Db 121 AGATGCACAGGAGGAACATGAGTGGTGTGCTGGGGGACCCCGACACCATCTTTC 180  
 QY 340 CAAGGTCCACCGTGATCAACATCCACAGCAGACCTCCGTGCCCGACCATGCTGTGGT 399  
 Db 181 CAAGGTCCACCGTGATCAACATCCACAGCAGACCTCCGTGCCCGACCATGCTGTGGT 240  
 QY 400 CCTGTTCAACACCTCTTCTTGAACCTGCTGTGGGCTTCATAGCATTCGCCTACT 459  
 Db 241 CCTGTTCAACACCTCTTCTTGAACCTGCTGTGGGCTTCATAGCATTCGCCTACT 300  
 QY 460 CCGTGAAGTCTAGGACAGGAAGATGGTTGGCGACGTGACCGGGGCCAGG 510  
 Db 301 CCGTGAAGTCTAGGACAGGAAGATGGTTGGCGACGTGACCGGGGCCAGG 351

RESULT 3

US-09-614-124B-1204  
 ; Sequence 1204, Application US/09614124B  
 ; Patent No. 6630574  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Wang, Tongtong  
 ; APPLICANT: Bangur, Chaitanya S.  
 ; APPLICANT: Lodes, Michael A.  
 ; APPLICANT: Fang, Gary  
 ; APPLICANT: Vedvick, Tom

; APPLICANT: Carter, Darrick  
 ; APPLICANT: Retter, Marc  
 ; APPLICANT: Mannion, Jane  
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND  
 ; FILE REFERENCE: 210121.478C9  
 ; CURRENT APPLICATION NUMBER: US/09/614.124B  
 ; CURRENT FILING DATE: 2001-07-11  
 ; NUMBER OF SEQ ID NOS: 1668  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 1204  
 ; LENGTH: 352  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapien  
 US-09-614-124B-1204

Query Match 31.7%; Score 351; DB 4; Length 352;  
 Best Local Similarity 100.0%; Pred. No. 1.3e-94;  
 Matches 351; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 160 GGGGAAGAGGAGTCTCACTGAGCACCGTCCAGCATCCGACACACAGCGGCCCTTCGC 219  
 Db 1 GGGGAAGAGGAGTCTCACTGAGCACCGTCCAGCATCCGACACACAGCGGCCCTTCGC 60  
 QY 220 TCCACGAGAAAACACACACTTCTCAAACTTCACTCAACACTTCTTCCCAAAGCCAGA 279  
 Db 61 TCCACGAGAAAACACACACTTCTCAAACTTCACTCAACACTTCTTCCCAAAGCCAGA 120  
 QY 280 AGATGCACAGGAGGAACATGAGTGGTGTGCTGGGGGACCCCGACACCATCTTTC 339  
 Db 121 AGATGCACAGGAGGAACATGAGTGGTGTGCTGGGGGACCCCGACACCATCTTTC 180  
 QY 340 CAAGGTCCACCGTGATCAACATCCACAGCAGACCTCCGTGCCCGACCATGCTGTGGT 399  
 Db 181 CAAGGTCCACCGTGATCAACATCCACAGCAGACCTCCGTGCCCGACCATGCTGTGGT 240  
 QY 400 CCTGTTCAACACCTCTTCTTGAACCTGCTGTGGGCTTCATAGCATTCGCCTACT 459  
 Db 241 CCTGTTCAACACCTCTTCTTGAACCTGCTGTGGGCTTCATAGCATTCGCCTACT 300  
 QY 460 CCGTGAAGTCTAGGACAGGAAGATGGTTGGCGACGTGACCGGGGCCAGG 510  
 Db 301 CCGTGAAGTCTAGGACAGGAAGATGGTTGGCGACGTGACCGGGGCCAGG 351

RESULT 4

US-09-671-325-1204  
 ; Sequence 1204, Application US/09671325  
 ; Patent No. 6667154  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Wang, Tongtong  
 ; APPLICANT: Bangur, Chaitanya S.  
 ; APPLICANT: Lodes, Michael A.  
 ; APPLICANT: Fang, Gary  
 ; APPLICANT: Vedvick, Tom  
 ; APPLICANT: Carter, Darrick  
 ; APPLICANT: Retter, Marc  
 ; APPLICANT: Mannion, Jane  
 ; APPLICANT: Fan, Liqun  
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
 ; FILE REFERENCE: 210121.478C12  
 ; CURRENT APPLICATION NUMBER: US/09/671.325  
 ; CURRENT FILING DATE: 2000-09-26  
 ; NUMBER OF SEQ ID NOS: 1825  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 1204  
 ; LENGTH: 352  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapien  
 US-09-671-325-1204

Query Match 31.7%; Score 351; DB 4; Length 352;

Best Local Similarity 100.0%; Pred. No. 1.3e-94;  
Matches 351; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 160 GGGGAAAGAGGCTCTACTGAGCACCCTCCAGCATCCGGACACACAGCGGCCCTTCGC 219  
Db 1 GGGGAAAGAGGCTCTACTGAGCACCCTCCAGCATCCGGACACACAGCGGCCCTTCGC 60  
QY 220 TCACGAGAAAACACACTTCTCAAACTTCACTCAAACTTCTTCCCAAGCCAGA 279  
Db 61 TCACGAGAAAACACACTTCTCAAACTTCACTCAAACTTCTTCCCAAGCCAGA 120  
QY 280 AGATGACAGAGGAGCAATGAGGTGGCTGTGGGGGACCCCGCAGCACCATCTTC 339  
Db 121 AGATGACAGAGGAGCAATGAGGTGGCTGTGGGGGACCCCGCAGCACCATCTTC 180  
QY 340 CAAAGTCCACCGTGATCAACATCCACAGCAGACCTCCGTGCGCCGACCATGTCGTGTGT 399  
Db 181 CAAAGTCCACCGTGATCAACATCCACAGCAGACCTCCGTGCGCCGACCATGTCGTGTGT 240  
QY 400 CCTGTTCACACCCCTCTTCTTGAACGTGTCTGTGGGCTTCATAGCAATCGCTACT 459  
Db 241 CCTGTTCACACCCCTCTTCTTGAACGTGTCTGTGGGCTTCATAGCAATCGCTACT 300  
QY 460 CCGTGAAGTCTAGGACAGGAGATGTTGCGAGCTGACCGGGGCCAGG 510  
Db 301 CCGTGAAGTCTAGGACAGGAGATGTTGCGAGCTGACCGGGGCCAGG 351

RESULT 5  
US-08-991-789A-283/c  
; Sequence 283, Application US/08991789A  
; Patent No. 6225054  
; GENERAL INFORMATION:  
; APPLICANT: Prudakis, Tony N.  
; Smith, John M.  
; Reed, Steven G.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
; TREATMENT AND DIAGNOSIS OF BREAST CANCER  
; NUMBER OF SEQUENCES: 292  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Seed IP Law Group  
; STREET: 701 Fifth Avenue, Suite 6300  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: USA  
; ZIP: 98104-7092  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/991,789A  
; FILING DATE: 11-Dec-1997  
; CLASSIFICATION: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Potter, Jane E. R.  
; REGISTRATION NUMBER: 33,332  
; REFERENCE/DOCKET NUMBER: 210121.419C3  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 622-4900  
; TELEFAX: (206) 682-6031  
; INFORMATION FOR SEQ ID NO: 283:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 764 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; SEQUENCE DESCRIPTION: SEQ ID NO: 283:  
US-08-991-789A-283  
Query Match 30.3%; Score 335.8; DB 3; Length 764;  
Best Local Similarity 92.9%; Pred. No. 6.6e-90;

Matches 341; Conservative 18; Mismatches 7; Indels 1; Gaps 1;  
QY 463 TGAAGTCTAGGACAGGAAGATGTTGGCGACGTGACCGGGGCCCGCCTATGCTCCA 522  
Db 367 TCAAGTCTAGGACAGGAAGATGTTGGCGACGTGACCGGGGCCCGCCTATGCTCCA 308  
QY 523 CCGCCAGTCCCTGAACATCTGGGCCCTGATTCTGGGCATCTCATGACCATTTGATTCA 582  
Db 307 CCGCCAGTCCCTGAASATSTGGGCCCTGATTSTGGGCATCTCTGACCTTTGATTCA 248  
QY 583 TCTGTCTACTGGTATTCGGCTCTGTGACAGTCTACCATATTATTTAGATTAATACAGG 642  
Db 247 TCTGTCTACTGGTATTCGGCTCTGTGACAGTCTACCATATTATTTAGATTAATACAGG 188  
QY 643 AAAACGGGCTTACTAGTAGCGGCCCATAGCTGCAACCTTTGCACTCCCACTGTGCAATG 702  
Db 187 AAAACGGGCTTACTAGTAGCGGCCCATAGCTGCAACCTTTGCACTCCCACTGTGCAATG 128  
QY 703 CTGGCCCTGACCGTGGGGCTGTGGCCCTGCCCCCTTGGTCCCTCCCTAGATACAGCA 762  
Db 127 STGGCCCTGACCGTGGGGCTGTGGCCCTGCCCCCTTGGTCCCTCCCTAGATACAGCA 68  
QY 763 GTTTATACCCACACACCTGTCTACACTGACATTCATTAAGTG-ACGTGCTTTGTGAAAAA 821  
Db 67 GTTTATACCCACACACCTGTCTACACTGATGTCTCATTTCAATAAGTGACGTGCTTTGCGAAAAA 8  
QY 822 AAAACAA 828  
Db 7 AAAAAAA 1

RESULT 6  
US-09-062-451-283/c  
; Sequence 283, Application US/09062451  
; Patent No. 6344550  
; GENERAL INFORMATION:  
; APPLICANT: Prudakis, Tony N.  
; APPLICANT: Smith, John M.  
; APPLICANT: Reed, Steven G.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
; TREATMENT AND DIAGNOSIS OF BREAST CANCER  
; NUMBER OF SEQUENCES: 297  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Seed and BERRY LLP  
; STREET: 6300 Columbia Center, 701 Fifth Avenue  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: USA  
; ZIP: 98104-7092  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/062,451  
; FILING DATE: 04-APR-1997  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Makl, David J.  
; REGISTRATION NUMBER: 31,392  
; REFERENCE/DOCKET NUMBER: 210121.419C2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 622-4900  
; TELEFAX: (206) 682-6031  
; INFORMATION FOR SEQ ID NO: 283:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 764 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-09-062-451-283

Query Match 30.3%; Score 335.8; DB 4; Length 764;  
 Best Local Similarity 92.9%; Pred. No. 6.6e-90;  
 Matches 341; Conservative 18; Mismatches 7; Indels 1; Gaps 1;

QY 463 TGAAGTCTAGGACAGAGAGATGTTGGCGACGTGACCGGGGCCCGAGGCTATGCTCCA 522  
 DB 367 TCAAGKGTAGGACAGAGAGAGATGTTGGCGACGTGACCGGGGCCCGAGGCTATGCTCCA 308  
 QY 523 CGGCAAGTGGCTGAACATCTGGGCGCTGATTCTGGGATCCTCATGACCAITGGATTCA 582  
 DB 307 CGGCAAGTGGCTGAASATSTGGGCGCTGATTSTGGGATCCTCTGTGACCTTTGGATTCA 248  
 QY 583 TCCTGTTACTGGTATTCGGCTCTGTGACAGTCTACATATTTATTTACAGATAATACAGG 642  
 DB 247 TCSGTDAYTGGWATTCGGCTCTGTGACAGTCTACATATTTATTTACAGATAATACAGG 188  
 QY 643 AAAAACCAGGTTACTAGTAGCGGCCCATAGCCTGCAACCTTTGCACTCCACTGTGCAATG 702  
 DB 187 AAAAACCAGGTTACTAGTAGCGGCCCATAGCCTGCAACCTTTGCACTCCACTGTGCAATG 128  
 QY 703 CTGCCCCGTGACCGTGGGGCTGTTGCCCTGCCCCCTGGTCTGCCCCCTAGATACAGCA 762  
 DB 127 STGCCCCGTGACCGTGGGGCTGTTGCCCTGCCCCCTGGTCTGCCCCCTAGATACAGCA 68  
 QY 763 GTTTATACCCACACACCTGTCTACAGTGCATTCATTAATTAAGTG-ACGTGCTTTGTAATA 821  
 DB 67 GTTTATACCCACACACCTGTCTACAGTGCATTCATTAATTAAGTGACGTGCTTGGCAATA 8  
 QY 822 AAAACAA 828  
 DB 7 AAAAAA 1

RESULT 7  
 US-09-289-198-283/c  
 ; Sequence 283, Application US/09289198  
 ; Patent No. 6586570  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Frudakis, Tony N.  
 ; APPLICANT: Reed, Steven G.  
 ; APPLICANT: Mishner, Lynda  
 ; APPLICANT: Reed, Steven G.  
 ; APPLICANT: Reed, Steven G.  
 ; APPLICANT: Reed, Steven G.  
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
 ; FILE REFERENCE: 210121.419C5  
 ; CURRENT APPLICATION NUMBER: US 09/289,198  
 ; CURRENT FILING DATE: 1999-04-09  
 ; EARLIER APPLICATION NUMBER: US 09/062,451  
 ; EARLIER FILING DATE: 1998-04-17  
 ; EARLIER APPLICATION NUMBER: US 08/991,789  
 ; EARLIER FILING DATE: 1997-12-11  
 ; EARLIER APPLICATION NUMBER: US 08/838,762  
 ; EARLIER FILING DATE: 1997-04-09  
 ; EARLIER APPLICATION NUMBER: PCT/US97/00485  
 ; EARLIER FILING DATE: 1997-01-10  
 ; EARLIER APPLICATION NUMBER: US 08/700,014  
 ; EARLIER FILING DATE: 1996-08-20  
 ; EARLIER APPLICATION NUMBER: US 08/585,392  
 ; EARLIER FILING DATE: 1996-01-01  
 ; NUMBER OF SEQ ID NOS: 312  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 283  
 ; LENGTH: 764  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapien  
 ; FEATURE:  
 ; NAME/KEY: misc\_feature  
 ; LOCATION: (1)...(764)  
 ; OTHER INFORMATION: n = A,T,C or G  
 US-09-289-198-283

Query Match 30.3%; Score 335.8; DB 4; Length 764;  
 Best Local Similarity 92.9%; Pred. No. 6.6e-90;

Matches 341; Conservative 18; Mismatches 7; Indels 1; Gaps 1;

QY 463 TGAAGTCTAGGACAGAGAGATGTTGGCGACGTGACCGGGGCCCGAGGCTATGCTCCA 522  
 DB 367 TCAAGKGTAGGACAGAGAGAGATGTTGGCGACGTGACCGGGGCCCGAGGCTATGCTCCA 308  
 QY 523 CGGCAAGTGGCTGAACATCTGGGCGCTGATTCTGGGATCCTCATGACCAITGGATTCA 582  
 DB 307 CGGCAAGTGGCTGAASATSTGGGCGCTGATTSTGGGATCCTCTGTGACCTTTGGATTCA 248  
 QY 583 TCCTGTTACTGGTATTCGGCTCTGTGACAGTCTACATATTTATTTACAGATAATACAGG 642  
 DB 247 TCSGTDAYTGGWATTCGGCTCTGTGACAGTCTACATATTTATTTACAGATAATACAGG 188  
 QY 643 AAAAACCAGGTTACTAGTAGCGGCCCATAGCCTGCAACCTTTGCACTCCACTGTGCAATG 702  
 DB 187 AAAAACCAGGTTACTAGTAGCGGCCCATAGCCTGCAACCTTTGCACTCCACTGTGCAATG 128  
 QY 703 CTGCCCCGTGACCGTGGGGCTGTTGCCCTGCCCCCTGGTCTGCCCCCTAGATACAGCA 762  
 DB 127 STGCCCCGTGACCGTGGGGCTGTTGCCCTGCCCCCTGGTCTGCCCCCTAGATACAGCA 68  
 QY 763 GTTTATACCCACACACCTGTCTACAGTGCATTCATTAATTAAGTG-ACGTGCTTTGTAATA 821  
 DB 67 GTTTATACCCACACACCTGTCTACAGTGCATTCATTAATTAAGTGACGTGCTTGGCAATA 8  
 QY 822 AAAACAA 828  
 DB 7 AAAAAA 1

RESULT 8  
 US-09-429-755-283/c  
 ; Sequence 283, Application US/09429755A  
 ; Patent No. 6656480  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Frudakis, Tony N.  
 ; APPLICANT: Smith, John W.  
 ; APPLICANT: Reed, Steven G.  
 ; APPLICANT: Mishner, Lynda  
 ; APPLICANT: Reed, Steven G.  
 ; APPLICANT: Reed, Steven G.  
 ; APPLICANT: Reed, Steven G.  
 ; TITLE OF INVENTION: TREATMENT AND DIAGNOSIS OF BREAST CANCER  
 ; FILE REFERENCE: 210121.419C6  
 ; CURRENT APPLICATION NUMBER: US 09/429,755A  
 ; CURRENT FILING DATE: 1999-10-28  
 ; NUMBER OF SEQ ID NOS: 315  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 283  
 ; LENGTH: 764  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapien  
 ; FEATURE:  
 ; NAME/KEY: misc\_feature  
 ; LOCATION: (1)...(764)  
 ; OTHER INFORMATION: n = A,T,C or G  
 US-09-429-755-283

Query Match 30.3%; Score 335.8; DB 4; Length 764;  
 Best Local Similarity 92.9%; Pred. No. 6.6e-90;  
 Matches 341; Conservative 18; Mismatches 7; Indels 1; Gaps 1;

QY 463 TGAAGTCTAGGACAGAGAGATGTTGGCGACGTGACCGGGGCCCGAGGCTATGCTCCA 522  
 DB 367 TCAAGKGTAGGACAGAGAGAGATGTTGGCGACGTGACCGGGGCCCGAGGCTATGCTCCA 308  
 QY 523 CGGCAAGTGGCTGAACATCTGGGCGCTGATTCTGGGATCCTCATGACCAITGGATTCA 582  
 DB 307 CGGCAAGTGGCTGAASATSTGGGCGCTGATTSTGGGATCCTCTGTGACCTTTGGATTCA 248  
 QY 583 TCCTGTTACTGGTATTCGGCTCTGTGACAGTCTACATATTTATTTACAGATAATACAGG 642

Db 247 TCSTGTDAYTGGWATTCGGCTCTGTGACAGTCTACCATATTATGTTACAGATAATACAGG 188  
 QY 643 AAAAAAGGGGTTACTAGTAGCGGCCCATAGCTTGCACACTTTGCACTCCACTGTGCAATG 702  
 Db 187 AAAAAAGGGGTTABTAGTAGCGGCCCATAGCTTGCACACTTTGCACTCCACTGTGCAATG 128  
 QY 703 CTGGCCCTGACCGTGGGCTGTGTCGCCCTGCGCCCTTGGCTCCCTCCCTAGATACAGCA 762  
 Db 127 STGGCCCTGACCGTGGGCTGTGTCGCCCTGCGCCCTTGGCTCCCTCCCTAGATACAGCA 68  
 QY 763 GTTTATACCCACACACTGTCTACACTGACATTCATTAAGTG-ACGTGCTTGTGAAAAA 821  
 Db 67 GTTTATACCCACACACTGTCTACAGTGTCTATTCAATTAAGTGACAGTGTGCTTGGCGAAAA 8  
 QY 822 AAAAAA 828  
 Db 7 AAAAAA 1

RESULT 9  
 US-09-023-655-1279  
 ; Sequence 1279, Application US/09023655  
 ; Patent No. 6607879  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Cocks, Benjamin G.  
 ; APPLICANT: Susan G. Stuart  
 ; APPLICANT: Jeffrey J. Seilhamer  
 ; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE  
 ; TITLE OF INVENTION: EXPRESSION  
 ; NUMBER OF SEQUENCES: 1508  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESS: INCYTE PHARMACEUTICALS, INC.  
 ; STREET: 3174 PORTER DRIVE  
 ; CITY: PALO ALTO  
 ; STATE: CALIFORNIA  
 ; COUNTRY: USA  
 ; ZIP: 94304  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
 ; CURRENT APPLICATION DATA:  
 ; FILING DATE: HEREWITH  
 ; CLASSIFICATION:  
 ; PRIORITY APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/023.655  
 ; FILING DATE:  
 ; CLASSIFICATION:  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Zeller, Karen J.  
 ; REGISTRATION NUMBER: 37,071  
 ; REFERENCE/DOCKET NUMBER: PA-0001 US  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (650) 855-0555  
 ; TELEFAX: (650) 845-4166  
 ; INFORMATION FOR SEQ ID NO: 1279:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 808 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; IMMEDIATE SOURCE:  
 ; LIBRARY: GENBANK  
 ; CLONE: 9311374  
 ; US-09-023-655-1279

Query Match 25.4%; Score 281.8; DB 4; Length 808;  
 Best Local Similarity 80.2%; Pred. No. 7.6e-74;  
 Matches 384; Conservative 0; Mismatches 82; Indels 13; Gaps 4;  
 QY 128 GAAATACACACTTCTAGAGAACTGAACACAGAGGGGAAA-GGAGGTCTCACTGAG-CACC 185

Db 138 GAAAAAGAACTGTGTGAAACCCGAAACTATCTGGGGAAGGGAGGCTCACTGAGTAACC 197  
 QY 186 GTCCAGATCCGACACACAGCGGCGCTTTCGCTCCAGCGCAAGAAACACACTTCTCAA 245  
 Db 198 ATCCAGTAACCG--ACCGCGCTGTGTTTCGCTGGACACCATGAGTCACACTGTCCAA 255  
 QY 246 ACCTTCACTCAACACTT-----CCTTCCCAAGCAGACAGATGCAAGAGGAGAA 296  
 Db 256 ACCTTCTTCTCTCTGTCAACAGTGGCCAGCGCCCCCAACTATGAGATGCTCAAGGAGGAG 315  
 QY 297 CATGAGGTGGTGTGCTGGGGGACACCCCGCAGCACCCTTCCAAAGTCCACCGTCAATC 356  
 Db 316 CAGAGGTGTGCTGTGGGGGGGCCCCCAACCCCTGCTCCCCCGACGTCACCGTGATC 375  
 QY 357 AACATCCACAGCAGACCTCCGTGCGCGACCATGTCGTCTGCTGCTTCTTCAACACCTC 416  
 Db 376 CACATCCGACGAGACCTCCGTGCGCGACCATGTCGTCTGCTGCTTCTTCAACACCTC 435  
 QY 417 TTCTTGAACCTGCTGCTGGGCTTTCATAGCATTCGCTACTCTCGTGAAGTCTAGGGAC 476  
 Db 436 TTCAAGAACCCCTGCTGCTGGGCTTTCATAGCATTCGCTACTCGTGAAGTCTAGGGAC 495  
 QY 477 AGGAAGATGTTGGCGACGTGACCGGGGCCAGGCTATGCTCCACCCCAAGTGCCTG 536  
 Db 496 AGGAAGATGTTGGCGACGTGACCGGGGCCAGGCTATGCTCCACCCCAAGTGCCTG 555  
 QY 537 AACATCTGGGCGCTGATTCCTGGGCATCTCATGACCATGATTCATCTGTTACTGCT 595  
 Db 556 AACATCTGGGCGCTGATTCCTGGGCATCTCATGACCATCTCTCATCTATCTCATCCAGT 614

RESULT 10  
 US-09-370-838-119  
 ; Sequence 119, Application US/09370838  
 ; Patent No. 8444425  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Reed, Steven G.  
 ; APPLICANT: Lodes, Michael J.  
 ; APPLICANT: Mohamath, Roadoh  
 ; APPLICANT: Secrist, Heather  
 ; TITLE OF INVENTION: COMPOUNDS FOR THERAPY AND DIAGNOSIS OF  
 ; TITLE OF INVENTION: LUNG CANCER AND METHODS FOR THEIR USE  
 ; FILE REFERENCE: 210121.475C1  
 ; CURRENT APPLICATION NUMBER: US/09/370.838  
 ; CURRENT FILING DATE: 1999-08-09  
 ; EARLIER APPLICATION NUMBER: US 09/285.323  
 ; EARLIER FILING DATE: 1999-04-02  
 ; NUMBER OF SEQ ID NOS: 289  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 119  
 ; LENGTH: 642  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapien  
 ; US-09-370-838-119

Query Match 24.8%; Score 275.4; DB 4; Length 642;  
 Best Local Similarity 70.0%; Pred. No. 5.4e-72;  
 Matches 450; Conservative 0; Mismatches 146; Indels 47; Gaps 4;  
 QY 202 ACCACAGCGGCCCTTTCGCTCCACGAGAAACACACTTCTCAAACTTCACTCAACACT 261  
 Db 25 ACCGCGCTGTGCTTGGCTGGACACCATGATCACAATCTCCAAACCTTCTTCTCTCTG 84  
 QY 262 T-----CCTTCCCCAAGCCAGAGATGCAAGAGGAGAACATGAGTGGCTGTGC 312  
 Db 85 TCAACAGTGGCCAGCCCCCACTATGAGATGCTCAAGGAGGAGCAGAGTGGCTGTGC 144  
 QY 313 TGGGGGACACCCCGCAGCACCCTTCTTCAAGGTCCACCGTGTATCAATCCACAGCGAGA 372  
 Db 145 TGGGGGCGCCCAACAACCCCTGCTCCCGGACCTCCACCGTATCCACATCCGACGAGA 204  
 QY 373 CCTCGTGGCCGACCATGCTGCTGCTCTCTTCAACACCCCTCTTCTTGAATGGTGTCT 432

```

Db 205 CCTCCGTCGCCGACATCGTCTGGTCCCTCTTCAACACCTCTTCATGAACCCCTGCT 264
Qy 433 GTCTGGGCTTCATAGCATTGCTCTACTCGTGAAGTCTAGGACAGAAAGATGGTGGG 492
Db 265 GCCTGGGCTTCATAGCATTGCTCTACTCGTGAAGTCTAGGACAGAAAGATGGTGGG 324
Qy 493 ACGTGACGGGGCCAGCCTATGCTCCACCGCAAGTGCCTGAACATCTGGGCCCTGA 552
Db 325 ACGTGACGGGGCCAGCCTATGCTCCACCGCAAGTGCCTGAACATCTGGGCCCTGA 384
Qy 553 TTCTGGGCATCTCATGACCATTTGATTGATTCATCTGTTACTGTTATTCGGCTCTGTGACAG 612
Db 385 TTCTGGGCATCTCATGACCATTTGATTGATTCATCTGTTACTGTTATTCGGCTCTGTGACAG 444
Qy 613 TCTACCATTTATGTTACAGATAATACAGAAACGGGGTACTAGTAGCCGCCCATAG 672
Db 445 ATGGATGATCAGG-----AGGCATCACTGAGGCCAGGAGCTCTGCCCATGA 491
Qy 673 CCTCAACCTTTGCACTCCACTGTGCAATGCTGGCCCTGCAAGCTGGGGCTGTGGCCCT 732
Db 492 CCTGTATCCACGACTCTCAACTTCATTCCTCGCCCTG----- 530
Qy 733 GCQCCCTGTGCTGCCCTAGATACAGAGTTTATATACCAACACACTGCTGTACACTGAC 792
Db 531 ---CCCGGAGCCGAGTCTGTATCAGCCCTTTATCTCACACGCTTTTCTACAATGGC 587
Qy 793 ATTCATTAAGTG-ACGTGCTGTGTAATAAATAAATAAATAA 834
Db 588 ATTCATTAAGTGACGTTTCTGTGTAATAAATAAATAAATAA 630

```

RESULT 11

US-09-702-705-998  
 ; Sequence 998, Application US/09702705  
 ; Patent No. 6504810

GENERAL INFORMATION:

APPLICANT: Wang, Tongtong  
 APPLICANT: Bangur, Chaitanya S.  
 APPLICANT: Lodes, Michael A.

APPLICANT: Fanger, Gary  
 APPLICANT: Vedvick, Tom  
 APPLICANT: Carter, Darrick  
 APPLICANT: Retter, Marc  
 APPLICANT: Mannion, Jane  
 APPLICANT: Fan, Liqun

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
 TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER

FILE REFERENCE: 210121.478C14  
 CURRENT APPLICATION NUMBER: US/09/702,705  
 CURRENT FILING DATE: 2000-10-30

NUMBER OF SEQ ID NOS: 1833  
 SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 998  
 LENGTH: 207

TYPE: DNA  
 ORGANISM: Homo sapien

FEATURE:  
 NAME/KEY: misc\_feature  
 LOCATION: (1)...(207)

OTHER INFORMATION: n = A,T,C or G

US-09-702-705-998

Query Match 15.8%; Score 175.6; DB 4; Length 207;  
 Best Local Similarity 90.3%; Pred. No. 1.3e-42;  
 Matches 187; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

Qy 302 GGTGCTGTGCTGGGGGACCCCGGACCACTCTTCCAAAGTCCACCGTGATCAACAT 361  
 Db 1 GGTGCTGTGCTGGGGGCGGCGCCACAACTGTCTCCCGGACGTCACCGTGATCCACAT 60

Qy 362 CCACAGCAGACCTCGTGGCCCGGACCACTGTGGTCTGCTCTGTTCAACACCCCTCTTCT 421

```

Db 61 NCGCAGCAGACCTCGTGGTCCCGGACCATGTGCTGGTCCCTGTTCAACACCCCTCTTCTCAT 120
Qy 422 GAATGCTGTGCTCTGGGCTTCATAGCAATCGCTACTCGTGAAGTCTAGGACAGGAA 481
Db 121 GAACCCCTGCTGCTGGGCTTCATAGCAATCGCTACTCGTGAAGTCTAGGACAGGAA 180
Qy 482 GATGTTGGCAGCTGACCGGGGCCCA 508
Db 181 GATGTTGGCAGCTGACCGGGGCCCA 207

```

RESULT 12

US-09-736-457-998  
 ; Sequence 998, Application US/09736457  
 ; Patent No. 6509448

GENERAL INFORMATION:

APPLICANT: Wang, Tongtong  
 APPLICANT: Bangur, Chaitanya S.  
 APPLICANT: Lodes, Michael A.

APPLICANT: Fanger, Gary  
 APPLICANT: Vedvick, Tom  
 APPLICANT: Carter, Darrick  
 APPLICANT: Retter, Marc  
 APPLICANT: Mannion, Jane  
 APPLICANT: Fan, Liqun

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
 TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER

FILE REFERENCE: 210121.478C15  
 CURRENT APPLICATION NUMBER: US/09/736,457  
 CURRENT FILING DATE: 2000-12-13

NUMBER OF SEQ ID NOS: 1854  
 SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 998  
 LENGTH: 207

TYPE: DNA  
 ORGANISM: Homo sapien

FEATURE:  
 NAME/KEY: misc\_feature  
 LOCATION: (1)...(207)

OTHER INFORMATION: n = A,T,C or G

US-09-736-457-998

Query Match 15.8%; Score 175.6; DB 4; Length 207;  
 Best Local Similarity 90.3%; Pred. No. 1.3e-42;  
 Matches 187; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

Qy 302 GGTGCTGTGCTGGGGGACCCCGGACCACTCTTCCAAAGTCCACCGTGATCAACAT 361  
 Db 1 GGTGCTGTGCTGGGGGCGGCGCCACAACTGTCTCCCGGACGTCACCGTGATCCACAT 60

Qy 362 CCACAGCAGACCTCGTGGCCCGGACCACTGTGGTCTGCTCTGTTCAACACCCCTCTTCT 421  
 Db 61 NCGCAGCAGACCTCGTGGCCCGGACCACTGTGGTCTGCTCTGTTCAACACCCCTCTTCTCAT 120

Qy 422 GAATGCTGTGCTCTGGGCTTCATAGCAATCGCTACTCGTGAAGTCTAGGACAGGAA 481  
 Db 121 GAACCCCTGCTGCTGGGCTTCATAGCAATCGCTACTCGTGAAGTCTAGGACAGGAA 180

Qy 482 GATGTTGGCAGCTGACCGGGGCCCA 508  
 Db 181 GATGTTGGCAGCTGACCGGGGCCCA 207

RESULT 13

US-09-614-124B-998  
 ; Sequence 998, Application US/09614124B  
 ; Patent No. 6630574

GENERAL INFORMATION:

APPLICANT: Wang, Tongtong  
 APPLICANT: Bangur, Chaitanya S.  
 APPLICANT: Lodes, Michael A.

APPLICANT: Fanger, Gary



APPLICANT: Vedvick, Tom  
 APPLICANT: Carter, Darick  
 APPLICANT: Retter, Marc  
 APPLICANT: Mannion, Jane  
 TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND  
 TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER  
 FILE REFERENCE: 210121.478C9  
 CURRENT APPLICATION NUMBER: US/09/614,124B  
 CURRENT FILING DATE: 2001-07-11  
 NUMBER OF SEQ ID NOS: 1668  
 SOFTWARE: FastSeq for Windows Version 3.0  
 SEQ ID NO 998  
 LENGTH: 207  
 TYPE: DNA  
 ORGANISM: Homo sapien  
 FEATURE:  
 NAME/KEY: misc.feature  
 LOCATION: (1)...(207)  
 OTHER INFORMATION: n = A,T,C or G  
 US-09-614-124B-998

Query Match 15.8%; Score 175.6; DB 4; Length 207;  
 Best Local Similarity 90.3%; Pred. No. 1.3e-42;  
 Matches 187; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

QY 302 GGTGGTGTGCTGGGGGACCCCGGACGACCATCTTCCAAAGGTCACCGGTGATCAACAT 361  
 DB 1 GGTGGTGTGCTGGGGGCGGCCACACACCTGCTCCCGGACGTCACCGGTGATCCACAT 60  
 QY 362 CCACAGGAGACCTCGTGGCGGACCATGCTGCTGGTCCCTGTTCAACACCCCTCTTCTT 421  
 DB 61 NCGCAGGAGACCTCGTGGCGGACCATGCTGCTGGTCCCTGTTCAACACCCCTCTTCTT 120  
 QY 422 GAACCTGCTGCTGGGCTTCATAGCATTCGCTACTCCGTTGAAGTCTAGGACAGGAA 481  
 DB 121 GAACCCCTGCTGCTGGGCTTCATAGCATTCGCTACTCCGTTGAAGTCTAGGACAGGAA 180  
 QY 482 GATGTTGGCGACGTGACCGGGGCCA 508  
 DB 181 GATGTTGGCGACGTGACCGGGGCCA 207

RESULT 14  
 US-09-671-325-998  
 Sequence 998, Application US/09671325  
 Patent No. 6667154  
 GENERAL INFORMATION:  
 APPLICANT: Wang, Tongtong  
 APPLICANT: Bangur, Chaitanya S.  
 APPLICANT: Lodes, Michael A.  
 APPLICANT: Fanger, Gary  
 APPLICANT: Vedvick, Tom  
 APPLICANT: Carter, Darick  
 APPLICANT: Retter, Marc  
 APPLICANT: Mannion, Jane  
 APPLICANT: Fan, Liqun  
 TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
 TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER  
 FILE REFERENCE: 210121.478C12  
 CURRENT APPLICATION NUMBER: US/09/671,325  
 CURRENT FILING DATE: 2000-09-26  
 NUMBER OF SEQ ID NOS: 1825  
 SOFTWARE: FastSeq for Windows Version 3.0  
 SEQ ID NO 998  
 LENGTH: 207  
 TYPE: DNA  
 ORGANISM: Homo sapien  
 FEATURE:  
 NAME/KEY: misc.feature  
 LOCATION: (1)...(207)  
 OTHER INFORMATION: n = A,T,C or G  
 US-09-671-325-998

Query Match 6.8%; Score 75; DB 4; Length 907;  
 Best Local Similarity 58.6%; Pred. No. 2.4e-12;  
 Matches 130; Conservative 0; Mismatches 92; Indels 0; Gaps 0;

QY 372 ACCTCCGTGCGCGACCATGCTGCTGGTCCCTGTTCAACACCCCTCTTCTTGAACGTGTC 431

Query Match 15.8%; Score 175.6; DB 4; Length 207;  
 Best Local Similarity 90.3%; Pred. No. 1.3e-42;  
 Matches 187; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

QY 302 GGTGGTGTGCTGGGGGACCCCGGACGACCATCTTCCAAAGGTCACCGGTGATCAACAT 361  
 DB 1 GGTGGTGTGCTGGGGGCGGCCACACACCTGCTCCCGGACGTCACCGGTGATCCACAT 60  
 QY 362 CCACAGGAGACCTCGTGGCGGACCATGCTGCTGGTCCCTGTTCAACACCCCTCTTCTT 421  
 DB 61 NCGCAGGAGACCTCGTGGCGGACCATGCTGCTGGTCCCTGTTCAACACCCCTCTTCTT 120  
 QY 422 GAACCTGCTGCTGGGCTTCATAGCATTCGCTACTCCGTTGAAGTCTAGGACAGGAA 481  
 DB 121 GAACCCCTGCTGCTGGGCTTCATAGCATTCGCTACTCCGTTGAAGTCTAGGACAGGAA 180  
 QY 482 GATGTTGGCGACGTGACCGGGGCCA 508  
 DB 181 GATGTTGGCGACGTGACCGGGGCCA 207

RESULT 15  
 US-09-620-312D-652  
 Sequence 652, Application US/09620312D  
 Patent No. 6569662  
 GENERAL INFORMATION:  
 APPLICANT: Tang, Y. Tom  
 APPLICANT: Liu, Chenghua  
 APPLICANT: Asundi, Vinod  
 APPLICANT: Zhang, Jie  
 APPLICANT: Ren, Feiyan  
 APPLICANT: Chen, Rui-hong  
 APPLICANT: Zhao, Qing A.  
 APPLICANT: Wehrman, Tom  
 APPLICANT: Xue, Aidong J.  
 APPLICANT: Yang, Yonghong  
 APPLICANT: Wang, Jian-Rui  
 APPLICANT: Zhou, Ping  
 APPLICANT: Ma, Yungqing  
 APPLICANT: Wang, Dunrui  
 APPLICANT: Wang, Zhiwei  
 APPLICANT: John Tillinghast  
 APPLICANT: Dmanac, Radoje T.  
 TITLE OF INVENTION: No. 6569662el Nucleic Acids and  
 TITLE OF INVENTION: Polypeptides  
 FILE REFERENCE: 784CIP2B  
 CURRENT APPLICATION NUMBER: US/09/620,312D  
 CURRENT FILING DATE: 2000-07-19  
 PRIOR APPLICATION NUMBER: 09/552,317  
 PRIOR FILING DATE: 2000-04-25  
 PRIOR APPLICATION NUMBER: 09/488,725  
 PRIOR FILING DATE: 2000-01-21  
 NUMBER OF SEQ ID NOS: 1105  
 SOFTWARE: pt\_FL\_genes Version 1.0  
 SEQ ID NO 652  
 LENGTH: 907  
 TYPE: DNA  
 ORGANISM: Homo sapiens  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: (205)..(597)  
 FEATURE:  
 NAME/KEY: misc.feature  
 LOCATION: (1)..(907)  
 OTHER INFORMATION: n = a,t,c or g  
 US-09-620-312D-652

Query Match 6.8%; Score 75; DB 4; Length 907;  
 Best Local Similarity 58.6%; Pred. No. 2.4e-12;  
 Matches 130; Conservative 0; Mismatches 92; Indels 0; Gaps 0;

QY 372 ACCTCCGTGCGCGACCATGCTGCTGGTCCCTGTTCAACACCCCTCTTCTTGAACGTGTC 431

```

Db 352 ACCGAGTGAACGACTATTACCTGTGGTCCATCTTCAACTTCGTCTACCTCAACTTCTGC 411
Qy 432 TGTCTGGGCTTTCATAGCATTCGCCTACTCCGTGAAAGTCTAGGACAGGAAGATGGTTGGC 491
Db 412 TGCCTGGGCTTTCATCGCCTTGGCCTTACTCCCTCAAAGTCCGAGACAAGAGCTTCTCAAT 471
Qy 492 GACGTGACCGGGGCCAGGCCCTATGCCTCCACCGCCAAAGTGCCTGAACATCTGGGCCCTG 551
Db 472 GACCTGAATGGAGCCGTGGAGATGCAAGACGGCCCGCTGTTCACATCACCAGTTCT 531
Qy 552 ATTCTGGGCATCCTCATGACCATTGGATTTCATCTGTTACTG 593
Db 532 GCCCTGGCAGCCTCCTGCATCATCCTCGTCTTCACTTCTCTG 573

```

Search completed: March 11, 2004, 14:41:57  
 Job time : 52.6434 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 11, 2004, 03:37:21 ; Search time 258.418 Seconds  
(without alignments)  
15799.545 Million cell updates/sec

Title: US-10-084-817-37  
Perfect score: 1109  
Sequence: 1 cctgcaccaggagacactgg.....acanaattccaccctatccg 1109

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2432557 seqs, 1840798884 residues

Total number of hits satisfying chosen parameters: 4965114

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA:\*

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*
- 2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*
- 3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*
- 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq.\*
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*
- 6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq.\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq.\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq.\*
- 11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq.\*
- 12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*
- 13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq.\*
- 14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq.\*
- 15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq.\*
- 16: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*
- 17: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*
- 18: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1051	94.8	1109	14	US-10-093-766-3
2	1051	94.8	1109	14	US-10-084-817-37
3	1051	94.8	1109	14	US-10-051-835-5
4	703	63.4	1617	14	US-10-247-671-89
5	699.8	63.1	1095	14	US-10-240-965-263
6	696.2	62.8	842	9	US-09-920-300A-1743
7	696.2	62.8	842	13	US-10-033-528-1743
8	696.2	62.8	842	14	US-10-099-926-1743
9	696.2	62.8	853	9	US-09-880-107-2108
10	696.2	62.8	853	10	US-09-960-706-826
11	696.2	62.8	853	10	US-09-873-319-392
12	696.2	62.8	853	14	US-10-171-581-195
13	696.6	61.9	1421	10	US-09-814-353-19857
14	685.8	61.8	852	14	US-10-176-847-27
15	624.4	56.3	647	14	US-10-301-822-80

16	624.4	56.3	647	15	US-10-133-937-25	Sequence 25, Appl
17	624.4	56.3	647	15	US-10-159-563-25	Sequence 25, Appl
18	465	41.9	468	12	US-10-085-783A-39396	Sequence 39396, A
19	465	41.9	468	15	US-10-242-535A-39396	Sequence 39396, A
20	456.4	41.2	470	12	US-10-085-783A-42090	Sequence 42090, A
21	456.4	41.2	470	15	US-10-242-535A-42090	Sequence 42090, A
22	454.8	41.0	701	15	US-10-264-049-1413	Sequence 1413, Ap
23	433	39.0	432	10	US-09-918-995-19602	Sequence 19602, A
24	427.6	38.6	434	10	US-09-918-995-35636	Sequence 35636, A
25	411	37.1	411	12	US-10-085-783A-32308	Sequence 32308, A
26	411	37.1	411	15	US-10-242-535A-32308	Sequence 32308, A
27	376.4	33.9	378	14	US-10-043-487-111	Sequence 111, App
28	353.2	31.8	507	14	US-10-106-698-2527	Sequence 2527, Ap
29	351	31.7	352	9	US-09-736-457-1204	Sequence 1204, Ap
30	351	31.7	352	9	US-09-902-941-1204	Sequence 1204, Ap
31	351	31.7	352	9	US-09-849-626-1204	Sequence 1204, Ap
32	351	31.7	352	14	US-10-017-754-1204	Sequence 1204, Ap
33	351	31.7	352	14	US-10-113-872-1204	Sequence 1204, Ap
34	350	31.6	350	9	US-09-796-692-5601	Sequence 5601, Ap
35	350	31.6	350	14	US-10-040-862-5601	Sequence 5601, Ap
36	350	31.6	350	15	US-10-057-475B-5601	Sequence 5601, Ap
37	350	31.6	350	15	US-10-154-884B-5601	Sequence 5601, Ap
38	340.8	30.7	622	14	US-10-299-497-2	Sequence 2, Appl
39	337.8	30.4	427	10	US-09-814-353-14867	Sequence 14867, A
40	335.8	30.3	764	9	US-09-810-936-383	Sequence 283, App
41	335.8	30.3	764	9	US-09-429-755-283	Sequence 283, App
42	335.8	30.3	764	9	US-09-924-400-283	Sequence 283, App
43	335.8	30.3	764	14	US-10-212-679-283	Sequence 283, App
44	332	29.9	341	9	US-09-796-692-3854	Sequence 3854, App
45	332	29.9	341	14	US-10-040-862-3854	Sequence 3854, Ap

ALIGNMENTS

RESULT 1  
US-10-093-766-3  
; Sequence 3, Application US/10093766  
; Publication No. US20030013099A1  
; GENERAL INFORMATION:  
; APPLICANT: Lasek, Amy W.  
; APPLICANT: Jones, David A.  
; APPLICANT: Karpf, Adam R.  
; TITLE OF INVENTION: GENES REGULATED BY DNA METHYLATION IN COLON TUMORS  
; FILE REFERENCE: PA-0047 US  
; CURRENT APPLICATION NUMBER: US/10/093,766  
; CURRENT FILING DATE: 2002-03-07  
; NUMBER OF SEQ ID NOS: 61  
; SOFTWARE: PERL Program  
; SEQ ID NO 3  
; LENGTH: 1109  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No. US20030013099A1 3094766CB1  
; NAME/KEY: unsure  
; LOCATION: 953, 963-996, 1002-1012, 1016-1018, 1022-1023, 1030, 1032-1034, 1053,  
; LOCATION: 1061, 1093  
; OTHER INFORMATION: a, t, c, g, or other  
; US-10-093-766-3

Query Match	94.8%	Score	1051	DB	14	Length	1109
Best Local Similarity	100.0%	Pred. No.	0				
Matches	1109	Conservative	0	Mismatches	0	Indels	0
Gaps	0						
QY	1	CCTGCACGAGACACTGGGAGGTTTGTAGTCCCAAAACCCGACACAGAGGAGCTGCAGC	60				
Db	1	CCTGCACGAGACACTGGGAGGTTTGTAGTCCCAAAACCCGACACAGAGGAGCTGCAGC	60				
QY	61	CTGAGGAAAGCAAGGATTTTCAGAGAGAGCGCTCGGCAAGTGTAGCAGCAATAGAAA	120				
Db	61	CTGAGGAAAGCAAGGATTTTCAGAGAGAGCGCTCGGCAAGTGTAGCAGCAATAGAAA	120				





GENERAL INFORMATION:  
APPLICANT: Mikita, Thomas  
APPLICANT: Shiffman, Dov  
APPLICANT: Porter, Gordon, J.  
APPLICANT: Kaser, Matthew R.  
TITLE OF INVENTION: GENES EXPRESSED IN TREATED FOAM CELLS  
FILE REFERENCE: PA-0050 US  
CURRENT APPLICATION NUMBER: US/10/247,671  
CURRENT FILING DATE: 2002-09-18  
PRIOR APPLICATION NUMBER: 60/323,784  
PRIOR FILING DATE: 2001-09-19  
NUMBER OF SEQ ID NOS: 186  
SOFTWARE: PERL Program  
SEQ ID NO 89  
LENGTH: 1617  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc feature  
OTHER INFORMATION: Incyte ID No. US20030194721A1 1328372.14  
US-10-247-671-89

Query Match 63.4%; Score 703; DB 14; Length 1617;  
Best Local Similarity 97.2%; Pred. No. 1.9e-210;  
Matches 726; Conservative 0; Mismatches 20; Indels 1; Gaps 1;

QY 106 AGCAGGAATAGAACTTAAGAGAAATACACACTTCTGAGAACTGAAACGACAGGGGAA 165  
DB 379 AGCAGGAATAGAACTTAAGAGAAATACACACTTCTGAGAACTGAAACGACAGGGGAA 438

QY 166 AGGAGGTCTACTGAGCAGCGTCCAGCATCCGGACACACAGCGGCGCTTGGCTCCACG 225  
DB 439 AGGAGGTCTACTGAGCAGCGTCCAGCATCCGGACACACAGCGGCGCTTGGCTCCACG 498

QY 226 CAGAAACCACTTCTCAAACTTCACTCAACCTTCTTCCCAAGCCAGCATGC 295  
DB 499 CAGAAACCACTTCTCAAACTTCACTCAACCTTCTTCCCAAGCCAGCATGC 558

QY 286 ACAAGGAGGAATGAGGTGGTGTGCTGGGGGACACCCCGGACCATCTTCCCAAGGT 345  
DB 559 ACAAGGAGGAATGAGGTGGTGTGCTGGGGGACACCCCGGACCATCTTCCCAAGGT 618

QY 346 CCACCGTGATCAACATCCACAGCAGACCTCCGTGCGCCGACCATGTCGTGCTGCTGCT 405  
DB 619 CCACCGTGATCAACATCCACAGCAGACCTCCGTGCGCCGACCATGTCGTGCTGCTGCT 678

QY 406 TCAACACCTCTTCTTGAACCTGGTGTGCTGGGCTTCATAGCATTCGCTACTCCGTTGA 465  
DB 679 TCAACACCTCTTCTTGAACCTGGTGTGCTGGGCTTCATAGCATTCGCTACTCCGTTGA 738

QY 466 AGTCTAGGAGCAGGAAGATGGTTGGGACGTGACCGGGGCGGCGCTATGCTCCACCG 525  
DB 739 AGTCTAGGAGCAGGAAGATGGTTGGGACGTGACCGGGGCGGCGCTATGCTCCACCG 798

QY 526 CCAAGTGCTGAACATCTGGGCGCTGATTCGGGATCCTCATGACATTTGGATTATCC 585  
DB 799 CCAAGTGCTGAACATCTGGGCGCTGATTCGGGATCCTCATGACATTTGGATTATCC 858

QY 586 TGTACTGGTATTCGGCTCTGTGACAGTCTACATATTATGTTACAGATTAATACAGAAA 645  
DB 859 TGTACTGGTATTCGGCTCTGTGACAGTCTACATATTATGTTACAGATTAATACAGAAA 918

QY 646 AACGGGTTACTAGTAGCGCGCCATAGCCTGCAACCTTTGCACTCCACTGTGCAATGCTG 705  
DB 919 AACGGGTTACTAGTAGCGCGCCATAGCCTGCAACCTTTGCACTCCACTGTGCAATGCTG 978

QY 706 GCCCTGACGCTGGGCTGTGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 765  
DB 979 GCCCTGACGCTGGGCTGTGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1038

QY 766 TATACCCACACCTGTCTACACTGACATTCATTAAGTG-AGTGTCTTGTGAAAAAA 824  
DB 1039 TATACCCACACCTGTCTACACTGACATTCATTAAGTGACAGTGTCTTGTGATGCT 1098

QY 825 ACAAATAAAACCCAGAGGGGGGGCGCG 851  
DB 1099 GTGACTTATCTGAGTGGGGGGCGCTG 1125

RESULT 5  
US-10-240-965-263  
Sequence 263, Application US/10240965  
Publication No. US20030165924A1  
GENERAL INFORMATION:  
APPLICANT: INCYTE GENOMICS, INC.  
APPLICANT: SHIFFMAN, Dov  
APPLICANT: SOMOGYI, Roland  
APPLICANT: LAM, Richard M.  
APPLICANT: SEILHAMER, Jeffrey J.  
APPLICANT: PORTER, Gordon J.  
APPLICANT: MIKITA, Thomas  
APPLICANT: TAI, Julie  
TITLE OF INVENTION: GENES EXPRESSED IN FOAM CELL DIFFERENTIATION  
FILE REFERENCE: PA-0025 PCT  
CURRENT APPLICATION NUMBER: US/10/240,965  
CURRENT FILING DATE: 2002-10-04  
PRIOR APPLICATION NUMBER: 60/195,106  
PRIOR FILING DATE: 2000-04-05  
NUMBER OF SEQ ID NOS: 276  
SOFTWARE: PERL Program  
SEQ ID NO 263  
LENGTH: 1095  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc feature  
OTHER INFORMATION: Incyte ID No. US20030165924A1 988231.7  
US-10-240-965-263

Query Match 63.1%; Score 699.8; DB 14; Length 1095;  
Best Local Similarity 99.6%; Pred. No. 1.6e-209;  
Matches 712; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 106 AGCAGGAATAGAACTTAAGAGAAATACACACTTCTGAGAACTGAAACGACAGGGGAA 165  
DB 381 AGCAGGAATAGAACTTAAGAGAAATACACACTTCTGAGAACTGAAACGACAGGGGAA 440

QY 166 AGGAGGTCTACTGAGCAGCGTCCCGACATCCGGACACACAGCGGCGCTTGGCTCCACG 225  
DB 441 AGGAGGTCTACTGAGCAGCGTCCCGACATCCGGACACACAGCGGCGCTTGGCTCCACG 500

QY 226 CAGAAACCACTTCTCAAACTTCACTCAACCTTCTTCCCAAGCCAGCATGC 285  
DB 501 CAGAAACCACTTCTCAAACTTCACTCAACCTTCTTCCCAAGCCAGCATGC 560

QY 286 ACAAGGAGGAATGAGGTGGTGTGCTGGGGGACACCCCGGACCATCTTCCCAAGGT 345  
DB 561 ACAAGGAGGAATGAGGTGGTGTGCTGGGGGACACCCCGGACCATCTTCCCAAGGT 620

QY 346 CCACCGTGATCAACATCCACAGCAGACCTCCGTGCGCCGACCATGTCGTGCTGCTGCTGCT 405  
DB 621 CCACCGTGATCAACATCCACAGCAGACCTCCGTGCGCCGACCATGTCGTGCTGCTGCTGCT 680

QY 406 TCAACACCTCTTCTTGAACCTGGTGTGCTGGGCTTCATAGCATTCGCTACTCCGTTGA 465  
DB 681 TCAACACCTCTTCTTGAACCTGGTGTGCTGGGCTTCATAGCATTCGCTACTCCGTTGA 740

QY 466 AGTCTAGGAGCAGGAAGATGGTTGGGACGTGACCGGGGCGGCGCTATGCTCCACCG 525  
DB 741 AGTCTAGGAGCAGGAAGATGGTTGGGACGTGACCGGGGCGGCGCTATGCTCCACCG 800

QY 526 CCAAGTGCTGAACATCTGGGCGCTGATTCGGGATTCCTCATGACATTTGGATTATCC 585  
DB 801 CCAAGTGCTGAACATCTGGGCGCTGATTCGGGATTCCTCATGACATTTGGATTATCC 860

QY 586 TGTACTGGTATTCGGCTCTGTGACAGTCTACCATATTATGTTACAGATAATACAGGAAA 645



Db 861 TGTACTGGTATTTCGGCTCTGTGACAGTCTACCATATTATTTACAGATAATACAGGAAA 920  
QY 646 AACGGGGTTACTAGTAGCCGCCCATAGCCTGCAACCTTTTGCACTCCACTGTGCAATGCTG 705  
Db 921 AACGGGGTTACTAGTAGCCGCCCATAGCCTGCAACCTTTTGCACTCCACTGTGCAATGCTG 980  
QY 706 GGCCTGCAGCTGGGGCTGTGGCCCTGCCCCCTTGGTCTGCCCCCTAGATACAGCAGTT 765  
Db 981 GGCCTGCAGCTGGGGCTGTGGCCCTGCCCCCTTGGTCTGCCCCCTAGATACAGCAGTT 1040  
QY 766 TATACCCACACACTGTCTACACTGACATTCAATAAAGTG-ACGTGCTTGTGAAA 819  
Db 1041 TATACCCACACACTGTCTACAGTGTCTATCAATAAAGTGACGTGCTTGTGAAA 1095

## RESULT 6

US-09-920-300A-1743  
; Sequence 1743, Application US/09920300A  
; Patent No. US20020136728A1  
; GENERAL INFORMATION:  
; APPLICANT: King, Gordon E.  
; APPLICANT: Meagher, Madeleine Joy  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Secrist, Heather  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER  
; FILE REFERENCE: 210121.547  
; CURRENT APPLICATION NUMBER: US/09/920.300A  
; CURRENT FILING DATE: 2001-07-31  
; NUMBER OF SEQ ID NOS: 1789  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1743  
; LENGTH: 842  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-920-300A-1743

Query Match 62.8%; Score 696.2; DB 9; Length 842;  
Best Local Similarity 99.4%; Pred. No. 1.9e-208;  
Matches 709; Conservative 0; Mismatches 3; Indels 1; Gaps 1;  
QY 106 AGCAGGAATAGAACTTAAGAGAAATACACACTTCTGAGAACTTGAACACGACAGGGAA 165  
Db 130 AGCAGGAATAGAACTTAAGAGAAATACACACTTCTGAGAACTTGAACACGACAGGGAA 189  
QY 166 AGGAGTCTCACTGAGCACCCTCCAGCATCCGGACACACAGCGGCCCTTCGCTCCAG 225  
Db 190 AGGAGTCTCACTGAGCACCCTCCAGCATCCGGACACACAGCGGCCCTTCGCTCCAG 249  
QY 226 CAGAAACCACTTCTCAACCTTCACTCAACACTTCTTCCCAAGCCAGAGATGC 285  
Db 250 CAGAAACCACTTCTCAACCTTCACTCAACACTTCTTCCCAAGCCAGAGATGC 309  
QY 286 ACAAGAGGAACATGAGTGGCTGTGCTGGGGGACACCCCGACACCATCTTCCAAGT 345  
Db 310 ACAAGAGGAACATGAGTGGCTGTGCTGGGGGACACCCCGACACCATCTTCCAAGT 369  
QY 346 CCACCGTGATCAATCCACAGCAGACCTCCGTGCCGACCATGTGCTGGTCCCTGT 405  
Db 370 CCACCGTGATCAATCCACAGCAGACCTCCGTGCCGACCATGTGCTGGTCCCTGT 429  
QY 406 TCAACACCCCTTCTTGAACCTGTGCTGCTGGGGCTTCATAGCATTCGGCTACTCCGTGA 465  
Db 430 TCAACACCCCTTCTTGAACCTGTGCTGCTGGGGCTTCATAGCATTCGGCTACTCCGTGA 489  
QY 466 AGTCTAGGACAGAAAGATGTTGGCGAGCTGACCGGGGCCAGGCCCTATGCTCCACCG 525  
Db 490 AGTCTAGGACAGAAAGATGTTGGCGAGCTGACCGGGGCCAGGCCCTATGCTCCACCG 549  
QY 526 CCAAGTGCCTGAACATCTTGGGCCCTTGATTTCTGGGCATCTCTATGACCATTTGATTCATCC 585  
Db 550 CCAAGTGCCTGAACATCTTGGGCCCTTGATTTCTGGGCATCTCTATGACCATTTGATTCATCC 609

QY 586 TGTACTGGTATTTCGGCTCTGTGACAGTCTACCATATTATTTACAGATAATACAGGAAA 645  
Db 610 TGTACTGGTATTTCGGCTCTGTGACAGTCTACCATATTATTTACAGATAATACAGGAAA 669  
QY 646 AACGGGGTTACTAGTAGCCGCCCATAGCCTGCAACCTTTTGCACTCCACTGTGCAATGCTG 705  
Db 670 AACGGGGTTACTAGTAGCCGCCCATAGCCTGCAACCTTTTGCACTCCACTGTGCAATGCTG 729  
QY 706 GGCCTGCAGCTGGGGCTGTGGCCCTGCCCCCTTGGTCTGCCCCCTAGATACAGCAGTT 765  
Db 730 GGCCTGCAGCTGGGGCTGTGGCCCTGCCCCCTTGGTCTGCCCCCTAGATACAGCAGTT 789  
QY 766 TATACCCACACACTGTCTACACTGACATTCAATAAAGTG-ACGTGCTTGTGAA 817  
Db 790 TATACCCACACACTGTCTACAGTGTCTATCAATAAAGTGACGTGCTTGTGAA 842

## RESULT 7

US-10-033-528-1743  
; Sequence 1743, Application US/10033528  
; Publication No. US20020131971A1  
; GENERAL INFORMATION:  
; APPLICANT: King, Gordon E.  
; APPLICANT: Meagher, Madeleine Joy  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Secrist, Heather  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER  
; FILE REFERENCE: 210121.547C1  
; CURRENT APPLICATION NUMBER: US/10/033.528  
; CURRENT FILING DATE: 2001-12-26  
; NUMBER OF SEQ ID NOS: 1896  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1743  
; LENGTH: 842  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-033-528-1743

Query Match 62.8%; Score 696.2; DB 13; Length 842;  
Best Local Similarity 99.4%; Pred. No. 1.9e-208;  
Matches 709; Conservative 0; Mismatches 3; Indels 1; Gaps 1;  
QY 106 AGCAGGAATAGAACTTAAGAGAAATACACACTTCTGAGAACTTGAACACGACAGGGAA 165  
Db 130 AGCAGGAATAGAACTTAAGAGAAATACACACTTCTGAGAACTTGAACACGACAGGGAA 189  
QY 166 AGGAGTCTCACTGAGCACCCTCCAGCATCCGGACACACAGCGGCCCTTCGCTCCAG 225  
Db 190 AGGAGTCTCACTGAGCACCCTCCAGCATCCGGACACACAGCGGCCCTTCGCTCCAG 249  
QY 226 CAGAAACCACTTCTCAACCTTCACTCAACACTTCTTCCCAAGCCAGAGATGC 285  
Db 250 CAGAAACCACTTCTCAACCTTCACTCAACACTTCTTCCCAAGCCAGAGATGC 309  
QY 286 ACAAGAGGAACATGAGTGGCTGTGCTGGGGGACACCCCGACACCATCTTCCAAGT 345  
Db 310 ACAAGAGGAACATGAGTGGCTGTGCTGGGGGACACCCCGACACCATCTTCCAAGT 369  
QY 346 CCACCGTGATCAATCCACAGCAGACCTCCGTGCCGACCATGTGCTGGTCCCTGT 405  
Db 370 CCACCGTGATCAATCCACAGCAGACCTCCGTGCCGACCATGTGCTGGTCCCTGT 429  
QY 406 TCAACACCCCTTCTTGAACCTGTGCTGCTGGGGCTTCATAGCATTCGGCTACTCCGTGA 465  
Db 430 TCAACACCCCTTCTTGAACCTGTGCTGCTGGGGCTTCATAGCATTCGGCTACTCCGTGA 489  
QY 466 AGTCTAGGACAGAAAGATGTTGGCGAGCTGACCGGGGCCAGGCCCTATGCTCCACCG 525  
Db 490 AGTCTAGGACAGAAAGATGTTGGCGAGCTGACCGGGGCCAGGCCCTATGCTCCACCG 549  
QY 526 CCAAGTGCCTGAACATCTTGGGCCCTTGATTTCTGGGCATCTCTATGACCATTTGATTCATCC 585

Db 550 CCAAGTGCCTGAAACATCTGGGCGCTGATTTCTGGGCATCCTCATGACCAATTGGATTTCATCC 609  
Qy 586 TGTACTGGTATTGGCTCTGTGACAGTCTACCATATTATTTACAGATAATACAGGAAA 645  
Db 610 TGTCACTGGTATTGGCTCTGTGACAGTCTACCATATTATTTACAGATAATACAGGAAA 669  
Qy 646 AACGGGGTTACTAGTAGCGGCCCATAGCTGCAACCTTTGCACTCCACTGTGCAATGCTG 705  
Db 670 AACGGGGTTACTAGTAGCGGCCCATAGCTGCAACCTTTGCACTCCACTGTGCAATGCTG 729  
Qy 706 GCGCTGACGCTGGGCTGTTGCCCTGCGCCCTTGGTCTGCTGCTGCTGCTGCTGCTGCTG 765  
Db 730 GCGCTGACGCTGGGCTGTTGCCCTGCGCCCTTGGTCTGCTGCTGCTGCTGCTGCTGCTG 789  
Qy 766 TATACCCACACACCTGTCTACACTGACATTCATTAAGTG-ACGTGCTTGTGA 817  
Db 790 TATACCCACACACCTGTCTACAGTGTCTCAATTAAGTGACAGTGTGCTTGTGA 842

RESULT 8  
US-10-099-926-1743  
; Sequence 1743, Application US/10099926  
; Publication No. US20030166064A1  
; GENERAL INFORMATION:  
; APPLICANT: King, Gordon E.  
; APPLICANT: Mesgher, Madeleine Joy  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Secrist, Heather  
; APPLICANT: Jiang, Yugu  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER  
; FILE REFERENCE: 210121.547C2  
; CURRENT APPLICATION NUMBER: US/10/099,926  
; CURRENT FILING DATE: 2002-03-17  
; NUMBER OF SEQ ID NOS: 1982  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1743  
; LENGTH: 842  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-099-926-1743

Query Match 62.8%; Score 696.2; DB 14; Length 842;  
Best Local Similarity 99.4%; Pred. No. 1.9e-208;  
Matches 709; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

Qy 106 AGCAGGAATAGAACTTAAGAGAAATACACACTTCTGAGAACTGAACGACAGGGGAA 165  
Db 130 AGCAGGAATAGAACTTAAGAGAAATACACACTTCTGAGAACTGAACGACAGGGGAA 189  
Qy 166 AGGAGGTCTCACTGAGCAGCGTCCAGCATCCGACACACAGCGGCCCTTGGTCCACG 225  
Db 190 AGGAGGTCTCACTGAGCAGCGTCCAGCATCCGACACACAGCGGCCCTTGGTCCACG 249  
Qy 226 CAGAAACACACACTTCTCAACCTTCACTCAACCTTCCCTCCCAAGCCAGAGATGC 285  
Db 250 CAGAAACACACACTTCTCAACCTTCACTCAACCTTCCCTCCCAAGCCAGAGATGC 309  
Qy 286 ACAAGGAGAAATAGAGTGGTGTGTGCTGGGGGACCCCGCAGCACCATCTTCCAAAGT 345  
Db 310 ACAAGGAGAAATAGAGTGGTGTGTGCTGGGGGACCCCGCAGCACCATCTTCCAAAGT 369  
Qy 346 CCACGCTGATCAACATCCACAGGAGACCTCGTGCCCGACCATGTGCTGGTCCCTGT 405  
Db 370 CCACGCTGATCAACATCCACAGGAGACCTCGTGCCCGACCATGTGCTGGTCCCTGT 429  
Qy 406 TCAACACCTCTTCTTGAACCTGTGTCTGGGCTTCATAGCATTCGCTTACTCCGTTGA 465  
Db 430 TCAACACCTCTTCTTGAACCTGTGTCTGGGCTTCATAGCATTCGCTTACTCCGTTGA 489  
Qy 466 AGTCTAGGACAGGAATAGTGGTGGAGCTGACCGGGGCCAGGCTATGCTCTCCACCG 525

Db 490 AGTCTAGGACAGGAAGATGTTGGCGAGTGAACCGGGGCCAGGGCTATGCTCTCCACCG 549  
Qy 526 CCAAGTGCCTGAAACATCTGGGCGCTGATTTCTGGGCATCCTCATGACCAATTGGATTTCATCC 585  
Db 550 CCAAGTGCCTGAAACATCTGGGCGCTGATTTCTGGGCATCCTCATGACCAATTGGATTTCATCC 609  
Qy 586 TGTACTGGTATTGGCTCTGTGACAGTCTACCATATTATTTACAGATAATACAGGAAA 645  
Db 610 TGTCACTGGTATTGGCTCTGTGACAGTCTACCATATTATTTACAGATAATACAGGAAA 669  
Qy 646 AACGGGGTTACTAGTAGCGGCCCATAGCTGCAACCTTTGCACTCCACTGTGCAATGCTG 705  
Db 670 AACGGGGTTACTAGTAGCGGCCCATAGCTGCAACCTTTGCACTCCACTGTGCAATGCTG 729  
Qy 706 GCGCTGACGCTGGGCTGTTGCCCTGCGCCCTTGGTCTGCTGCTGCTGCTGCTGCTGCTG 765  
Db 730 GCGCTGACGCTGGGCTGTTGCCCTGCGCCCTTGGTCTGCTGCTGCTGCTGCTGCTGCTG 789  
Qy 766 TATACCCACACACCTGTCTACACTGACATTCATTAAGTG-ACGTGCTTGTGA 817  
Db 790 TATACCCACACACCTGTCTACAGTGTCTCAATTAAGTGACAGTGTGCTTGTGA 842

RESULT 9  
US-09-880-107-2108  
; Sequence 2108, Application US/09880107  
; Patent No. US20020142981A1  
; GENERAL INFORMATION:  
; APPLICANT: Horie, Daisi T.  
; APPLICANT: Vockley, Joseph G.  
; APPLICANT: Scherf, Uwe  
; APPLICANT: Gene Logic, Inc.  
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer  
; FILE REFERENCE: 44921-5028-WO  
; CURRENT APPLICATION NUMBER: US/09/880,107  
; CURRENT FILING DATE: 2001-06-14  
; PRIOR APPLICATION NUMBER: US 60/211,379  
; PRIOR FILING DATE: 2000-06-14  
; PRIOR APPLICATION NUMBER: US 60/237,054  
; PRIOR FILING DATE: 2000-10-02  
; NUMBER OF SEQ ID NOS: 3950  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2108  
; LENGTH: 853  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 J04164  
US-09-880-107-2108

Query Match 62.8%; Score 696.2; DB 9; Length 853;  
Best Local Similarity 99.4%; Pred. No. 1.9e-208;  
Matches 709; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

Qy 106 AGCAGGAATAGAACTTAAGAGAAATACACACTTCTGAGAACTGAACGACAGGGGAA 165  
Db 141 AGCAGGAATAGAACTTAAGAGAAATACACACTTCTGAGAACTGAACGACAGGGGAA 200  
Qy 166 AGGAGGTCTCACTGAGCAGCGTCCAGCATCCGACACACAGCGGCCCTTGGTCCACG 225  
Db 201 AGGAGGTCTCACTGAGCAGCGTCCAGCATCCGACACACAGCGGCCCTTGGTCCACG 260  
Qy 226 CAGAAACACACACTTCTCAACCTTCACTCAACCTTCCCTCCCAAGCCAGAGATGC 285  
Db 261 CAGAAACACACACTTCTCAACCTTCACTCAACCTTCCCTCCCAAGCCAGAGATGC 320  
Qy 286 ACAAGGAGAAATAGAGTGGTGTGCTGGGGGACCCCGCAGCACCATCTTCCAAAGT 345  
Db 321 ACAAGGAGAAATAGAGTGGTGTGCTGGGGGACCCCGCAGCACCATCTTCCAAAGT 380  
Qy 346 CCACGCTGATCAACATCCACAGGAGACCTCGTGCCCGACCATGTGCTGGTCCCTGT 405  
Db 381 CCACGCTGATCAACATCCACAGGAGACCTCGTGCCCGACCATGTGCTGGTCCCTGT 440

QY 406 TCAACACCCCTCTTCTGAACTGGTGTCTGTGGCTTCATAGCAATCGCTACTCCGTGA 465  
 DB 441 TCAACACCCCTCTTCTGAACTGGTGTCTGTGGCTTCATAGCAATCGCTACTCCGTGA 500  
 QY 466 AGTCTAGGACAGAGAGATGGTTGGCGACGTGACCGGGGCCCGAGCCCTATGCTCCACCG 525  
 DB 501 AGTCTAGGACAGAGAGATGGTTGGCGACGTGACCGGGGCCCGAGCCCTATGCTCCACCG 560  
 QY 526 CCAAGTGCCTGAACATCTGGGCTGTGACAGTCTACCATATTTATGTTACAGATAATACAGGAAA 585  
 DB 561 CCAAGTGCCTGAACATCTGGGCTGTGACAGTCTACCATATTTATGTTACAGATAATACAGGAAA 620  
 QY 586 TGTACTGTGTAATCGGCTGTGACAGTCTACCATATTTATGTTACAGATAATACAGGAAA 645  
 DB 621 TGTACTGTGTAATCGGCTGTGACAGTCTACCATATTTATGTTACAGATAATACAGGAAA 680  
 QY 706 GCCTGCACGCTGGGCTGTGACAGTCTACCATATTTATGTTACAGATAATACAGGAAA 765  
 DB 741 GCCTGCACGCTGGGCTGTGACAGTCTACCATATTTATGTTACAGATAATACAGGAAA 800  
 QY 766 TATACCCACACACTGTCTACACTGACATTCATTAATAAGTG-ACGTGCTTGTGA 817  
 DB 801 TATACCCACACACTGTCTACACTGACATTCATTAATAAGTG-ACGTGCTTGTGA 853

RESULT 10  
 US-09-960-706-626  
 ; Sequence 626, Application US/09960706  
 ; Publication No. US20030134280A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Mungler, William E.  
 ; TITLE OF INVENTION: Identifying Drugs for and Diagnosis of Benign Prostatic Hyperplasia  
 ; FILE REFERENCE: 44921-5029-01US  
 ; CURRENT APPLICATION NUMBER: US/09/960,706  
 ; PRIOR FILING DATE: 2001-09-24  
 ; PRIOR APPLICATION NUMBER: 60/223,323  
 ; PRIOR FILING DATE: 2000-08-07  
 ; PRIOR APPLICATION NUMBER: 09/873,319  
 ; PRIOR FILING DATE: 2001-06-05  
 ; NUMBER OF SEQ ID NOS: 1124  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 626  
 ; LENGTH: 853  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; OTHER INFORMATION: Genbank Accession No. US20030134280A1 J04164  
 US-09-960-706-626

Query Match 62.8%; Score 696.2; DB 10; Length 853;  
 Best Local Similarity 99.4%; Pred. No. 1.9e-208;  
 Matches 709; Conservative 0; Mismatches 3; Indels 1; Gaps 1;  
 QY 106 AGCAGGAAATAGAACTTAAGAGAAATACACACTTCTGAGAACTGAAACGACAGGGGAA 165  
 DB 141 AGCAGGAAATAGAACTTAAGAGAAATACACACTTCTGAGAACTGAAACGACAGGGGAA 200  
 QY 166 AGAGGTCTCACTGAGCACCGTCCGATCCGACACACAGCGGCCCTTGGTCCACG 225  
 DB 201 AGAGGTCTCACTGAGCACCGTCCGATCCGACACACAGCGGCCCTTGGTCCACG 260  
 QY 226 CAGAAACACACACTTCTCAACCTTCACTCAACACTTCTTCCCAAGCCAGAGATGC 285  
 DB 261 CAGAAACACACACTTCTCAACCTTCACTCAACACTTCTTCCCAAGCCAGAGATGC 320  
 QY 286 AAGAGGAAATAGAGTGTGTGGGGGACCCCGGACACCATCTTCCAGGT 345

DB 321 ACAAGGAGGAACATGAGGTGCTGTGTGGGGCACCCCCCAGCAGCACCATCTTCCAAGT 380  
 QY 346 CCAACCGTGATCAACATCCACAGAGACCTCGGTGCCCGACCATCTGCTGGTCCCTGT 405  
 DB 381 CCAACCGTGATCAACATCCACAGAGACCTCGGTGCCCGACCATCTGCTGGTCCCTGT 440  
 QY 406 TCAACACCCCTCTTCTTGAACCTGGTGTCTGTGGGCTTCATAGCAATCGCTACTCCGTGA 465  
 DB 441 TCAACACCCCTCTTCTTGAACCTGGTGTCTGTGGGCTTCATAGCAATCGCTACTCCGTGA 500  
 QY 466 AGTCTAGGACAGAGAGATGGTTGGCGACGTGACCGGGGCCCGAGCCCTATGCTCCACCG 525  
 DB 501 AGTCTAGGACAGAGAGATGGTTGGCGACGTGACCGGGGCCCGAGCCCTATGCTCCACCG 560  
 QY 526 CCAAGTGCCTGAACATCTGGGCTGTGACAGTCTACCATATTTATGTTACAGATAATACAGGAAA 585  
 DB 561 CCAAGTGCCTGAACATCTGGGCTGTGACAGTCTACCATATTTATGTTACAGATAATACAGGAAA 620  
 QY 586 TGTACTGTGTAATCGGCTGTGACAGTCTACCATATTTATGTTACAGATAATACAGGAAA 645  
 DB 621 TGTACTGTGTAATCGGCTGTGACAGTCTACCATATTTATGTTACAGATAATACAGGAAA 680  
 QY 646 AACGGGTTACTAGTAGCCGCCCATAGCTGCAACCTTTGCACTCCACTGTGCAATGCTG 705  
 DB 681 AACGGGTTACTAGTAGCCGCCCATAGCTGCAACCTTTGCACTCCACTGTGCAATGCTG 740  
 QY 706 GCCTGCACGCTGGGCTGTGACAGTCTACCATATTTATGTTACAGATAATACAGGAAA 765  
 DB 741 GCCTGCACGCTGGGCTGTGACAGTCTACCATATTTATGTTACAGATAATACAGGAAA 800  
 QY 766 TATACCCACACACTGTCTACACTGACATTCATTAATAAGTG-ACGTGCTTGTGA 817  
 DB 801 TATACCCACACACTGTCTACACTGACATTCATTAATAAGTG-ACGTGCTTGTGA 853

RESULT 11  
 US-09-873-319-392  
 ; Sequence 392, Application US/09873319A  
 ; Publication No. US20030134324A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Mungler, William E.  
 ; APPLICANT: Kulkarni, Prakash  
 ; APPLICANT: Getzenberg, Robert H.  
 ; APPLICANT: Waga, Iwao  
 ; APPLICANT: Yamamoto, Jun  
 ; TITLE OF INVENTION: Identifying Drugs for and Diagnosis of Benign Prostatic  
 ; FILE REFERENCE: 44921-5029-US  
 ; CURRENT APPLICATION NUMBER: US/09/873,319A  
 ; CURRENT FILING DATE: 2001-06-05  
 ; EARLIER FILING DATE: 2000-08-07  
 ; NUMBER OF SEQ ID NOS: 755  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 392  
 ; LENGTH: 853  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; OTHER INFORMATION: Genbank Accession No. US20030134324A1 J04164  
 US-09-873-319-392

Query Match 62.8%; Score 696.2; DB 10; Length 853;  
 Best Local Similarity 99.4%; Pred. No. 1.9e-208;  
 Matches 709; Conservative 0; Mismatches 3; Indels 1; Gaps 1;  
 QY 106 AGCAGGAAATAGAACTTAAGAGAAATACACACTTCTGAGAACTGAAACGACAGGGGAA 165  
 DB 141 AGCAGGAAATAGAACTTAAGAGAAATACACACTTCTGAGAACTGAAACGACAGGGGAA 200  
 QY 166 AGAGGTCTCACTGAGCACCGTCCGATCCGACACACAGCGGCCCTTCCCTCCACG 225  
 DB 201 AGAGGTCTCACTGAGCACCGTCCGATCCGACACACAGCGGCCCTTCCCTCCACG 260

226	QY	CAGAAAACCACTTCTCAAAACCTTCACTCAACACTTCTCTCCCAAGGCGCAGAAGATGC	285
261	DB	CAGAAAACCACTTCTCTCAAAACCTTCACTCAACACTTCTCTCCCAAGGCGCAGAAGATGC	320
286	QY	ACAAGGAGGAAACATGAGTGGCTGTGCTGGGGGCACCCCGACGACCATCTCTCCAAAGT	345
321	DB	ACAAGGAGGAAACATGAGTGGCTGTGCTGGGGGCACCCCGACGACCATCTCTCCAAAGT	380
346	QY	CCACCGTGATCAACATCCACAGCGAGACCTCCCGTGGCGGACCATGTGCTCTGGTCCCTGT	405
381	DB	CCACCGTGATCAACATCCACAGCGAGACCTCCCGTGGCGGACCATGTGCTCTGGTCCCTGT	440
406	QY	TCAACACCCCTCTTCTTGAACCTGGTGTGCTGTGGCGTTTCATAGCATTTGGCTTACTCCGTGA	465
441	DB	TCAACACCCCTCTTCTTGAACCTGGTGTGCTGTGGCGTTTCATAGCATTTGGCTTACTCCGTGA	500
466	QY	AGTCTAGGGAACAGGAAGATGGTTGGCGACGTGACCGGGGCCAGGCGCTATGCTCTCAACG	525
501	DB	AGTCTAGGGAACAGGAAGATGGTTGGCGACGTGACCGGGGCCAGGCGCTATGCTCTCAACG	560
526	QY	CCAAGTGCCCTGAACATCTGGGCCCTGATTTCTGGGCATCCTCATGACCATTTGGATTCATCC	585
561	DB	CCAAGTGCCCTGAACATCTGGGCCCTGATTTCTGGGCATCCTCATGACCATTTGGATTCATCC	620
586	QY	TGTTACTGGTATTCCGGCTCTGTGACAGTCTACCATATTATGTTACAGNATAATACAGGAA	645
621	DB	TGTTACTGGTATTCCGGCTCTGTGACAGTCTACCATATTATGTTACAGNATAATACAGGAA	680
646	QY	AACGGGGTTACTAGTAGCCGCCCATAGCTCGCAACCTTTGCACCTCCACTGTGCAATGCTG	705
681	DB	AACGGGGTTACTAGTAGCCGCCCATAGCTCGCAACCTTTGCACCTCCACTGTGCAATGCTG	740
706	QY	GCCTTGACCGCTGGGCTGTGGCCCTGCGCCCTTGCTCTGGCTGCCCTAGATACAGCACTT	765
741	DB	GCCTTGACCGCTGGGCTGTGGCCCTGCGCCCTTGCTCTGGCTGCCCTAGATACAGCACTT	800
766	QY	TATACCCACACACCTGTCTCACTGCACATTCAAATAAGTG-ACGTGCTTGTGA	817
801	DB	TATACCCACACACCTGTCTCACTGCAGTGCATTCAAATAAGTGACACGTGCTTGTGA	853

RESULT 12

```

US-10-171-581-195
; Sequence 195, Application US/10171581
; Publication No. US20030104428A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: Linsley, Peter
; APPLICANT: Mao Mao
; TITLE OF INVENTION: Signature Genes in Chronic Myelogenous Leukemia
; FILE REFERENCE: 9301.157-999
; CURRENT APPLICATION NUMBER: US/10171.581
; CURRENT FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/298,914
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 366
; SEQ ID NO 195
; LENGTH: 853
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: J04164
; DATABASE ENTRY DATE: 2001-06-18
; US-10-171-581-195

```

Qy	166	AGAGGTCTCACTGAGCAACGTCCTCCAGCATCCGAGACCAACAGCGGCGCTTCGTCCACG	222
Db	201	AGAGGTCTCACTGAGCAACGTCCTCCAGCATCCGAGACCAACAGCGGCGCTTCGTCCACG	260
Qy	226	CAGAAAACCACTCTTCAAACTTCTCACTCAACACTTCTTCTTCCCAAAGCAGAAATGTC	285
Db	261	CAGAAAACCACTCTTCAAACTTCTCACTCAACACTTCTTCTTCCCAAAGCAGAAATGTC	320
Qy	286	ACAGGAGGAACATGAGTGCCTGTGTGGGGGACACCCCGAGACCACTCTTCAAAAGT	345
Db	321	ACAGGAGGAACATGAGTGCCTGTGTGGGGGACACCCCGAGACCACTCTTCAAAAGT	380
Qy	346	CCACCGTGATCAACATCCACAGCGAGACCTTCGTGCGCGGACCATGTCTGTGTCCTCTGT	405
Db	381	CCACCGTGATCAACATCCACAGCGAGACCTTCGTGCGCGGACCATGTCTGTGTCCTCTGT	440
Qy	406	TCAACACCCCTCTTCTTGAACCTGGTGCTGTCTGGGCTTCTATAGCATTTGCGCTACTCCGTGA	465
Db	441	TCAACACCCCTCTTCTTGAACCTGGTGCTGTCTGGGCTTCTATAGCATTTGCGCTACTCCGTGA	500
Qy	466	AGTCTAGGACAGGAAGATGGTTGGCGACGTGACCGGGGCGCCAGGCCTATGCCCTCCACCG	525
Db	501	AGTCTAGGACAGGAAGATGGTTGGCGACGTGACCGGGGCGCCAGGCCTATGCCCTCCACCG	560
Qy	526	CCAAGTGCCTGAAACATCTGGGCCCTGATTTCTGGGCATCTCTCATGACCAATGCGATTCATCC	585
Db	561	CCAAGTGCCTGAAACATCTGGGCCCTGATTTCTGGGCATCTCTCATGACCAATGCGATTCATCC	620
Qy	586	TGTTACTGTGTTATTCGGCTCTGTGACAGCTTACCATATTTATGTTACAGATAATACAGGAAA	645
Db	621	TGTTACTGTGTTATTCGGCTCTGTGACAGCTTACCATATTTATGTTACAGATAATACAGGAAA	680
Qy	646	AACGGGGTTACTAGTAGCGGCCCATAGCGCTGCAACCTTTGCACTCCACTGTGCAATGCTG	705
Db	681	AACGGGGTTACTAGTAGCGGCCCATAGCGCTGCAACCTTTGCACTCCACTGTGCAATGCTG	740
Qy	706	GCCCTGCACGCTGGGGCTGTTGCCCTGCCCTTGGTCTTCCCTTGGCTCCCTAGATACAGCAGTT	765
Db	741	GCCCTGCACGCTGGGGCTGTTGCCCTGCCCTTGGTCTTCCCTTGGCTCCCTAGATACAGCAGTT	800
Qy	766	TATACCCACACCTGTCTACACTGACATTCATTAAGTG-ACGTGCTGTGA	817
Db	801	TATACCCACACCTGTCTACACTGACATTCATTAAGTG-ACGTGCTGTGA	853
RESULT 13			
US-09-814-353-19857			
; Sequence 19857, Application US/09814353			
; Publication No. US20030165631A1			
; GENERAL INFORMATION:			
; APPLICANT: Lee, John			
; APPLICANT: Thompson, Pamela			
; APPLICANT: Lillie, James			
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR			
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND			
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER			
; FILE REFERENCE: MRI-006B			
; CURRENT APPLICATION NUMBER: US/09/814,353			
; CURRENT FILING DATE: 2001-03-21			
; PRIOR APPLICATION NUMBER: US 60/191,031			
; PRIOR FILING DATE: 2000-03-21			
; PRIOR APPLICATION NUMBER: US 60/207,124			
; PRIOR FILING DATE: 2000-05-25			
; PRIOR APPLICATION NUMBER: US 60/211,940			
; PRIOR FILING DATE: 2000-06-15			
; PRIOR APPLICATION NUMBER: US 60/216,820			
; PRIOR FILING DATE: 2000-07-07			
; PRIOR APPLICATION NUMBER: US 60/220,661			
; PRIOR FILING DATE: 2000-07-25			
; PRIOR APPLICATION NUMBER: US 60/257,672			
; PRIOR FILING DATE: 2000-12-21			
; NUMBER OF SEQ ID NOS: 22037			

RESIST. T 13

US-09-814-353-19857  
 ; Sequence 19857, Application US/09814353  
 ; Publication No. US20030165831A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Lee, John  
 ; APPLICANT: Thompson, Pamela  
 ; APPLICANT: Lillie, James  
 ; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR  
 ; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND  
 ; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER  
 ; FILE REFERENCE: MRI-006B  
 ; CURRENT APPLICATION NUMBER: US/09/814,353  
 ; CURRENT FILING DATE: 2001-03-21  
 ; PRIOR APPLICATION NUMBER: US 60/191,031  
 ; PRIOR FILING DATE: 2000-03-21  
 ; PRIOR APPLICATION NUMBER: US 60/207,124  
 ; PRIOR FILING DATE: 2000-05-25  
 ; PRIOR APPLICATION NUMBER: US 60/211,940  
 ; PRIOR FILING DATE: 2000-06-15  
 ; PRIOR APPLICATION NUMBER: US 60/216,820  
 ; PRIOR FILING DATE: 2000-07-07  
 ; PRIOR APPLICATION NUMBER: US 60/220,661  
 ; PRIOR FILING DATE: 2000-07-25  
 ; PRIOR APPLICATION NUMBER: US 60/257,672  
 ; PRIOR FILING DATE: 2000-12-21  
 ; NUMBER OF SEQ ID NOS: 22037

; SOFTWARE: PastSeq for Windows Version 4.0  
; SEQ ID NO 19857  
; LENGTH: 1421  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-814-353-19857

Query Match 61.9%; Score 586.6; DB 10; Length 1421;  
Best Local Similarity 98.6%; Pred. No. 2.6e-205; Indels 1; Gaps 1;  
Matches 703; Conservative 0; Mismatches 9;

QY 126 GAGAAATACACACTTCTGAGAAATCGAAACACAGGGAAGAGGTCTCACTGAGCAC 185  
DB 46 GAGAAATACACACTTCTGAGAAATCGAAACACAGGGAAGAGGTCTCACTGAGCAC 105

QY 186 GTCCAGCATCCGACACACAGGCGGCTTCGCTCCAGCAGAGAAACACACTTCTCAA 245  
DB 106 GTCCAGCATCCGACACACAGGCGGCTTCGCTCCAGCAGAGAAACACACTTCTCAA 165

QY 246 ACCTTCACCTCAACACTTCTTCCCAAGCCAGAGATGCAAGGAGGAAACATGAGGTG 305  
DB 166 ACCTTCACCTCAACACTTCTTCCCAAGCCAGAGATGCAAGGAGGAAACATGAGGTG 225

QY 306 GCTGTGCTGGGGGACCCCGACACATCTTCCAGGTCACCGTGATCAACATCCAC 365  
DB 226 GCTGTGCTGGGGGACCCCGACACATCTTCCAGGTCACCGTGATCAACATCCAC 285

QY 366 AGGAGACCTCCGTCGCGGACCATGTGCTGCTCCCTGTTCAACACCTCTTCTTGAAC 425  
DB 286 AGGAGACCTCCGTCGCGGACCATGTGCTGCTCCCTGTTCAACACCTCTTCTTGAAC 345

QY 426 TGGTGTGCTGGGCTTATAGCATTCGCTACTCCGTTAGGTCAGGAGGAGATG 485  
DB 346 TGGTGTGCTGGGCTTATAGCATTCGCTACTCCGTTAGGTCAGGAGGAGATG 405

QY 486 GTTGGGACGTGACCGGGGCGGCGCTATGCTCCACCGCAAGTCTGAAACATCTGG 545  
DB 406 GTTGGGACGTGACCGGGGCGGCGCTATGCTCCACCGCAAGTCTGAAACATCTGG 465

QY 546 GCCTGATTCGGGCTTCATGACATTCGATTCCTGTTACTGGTATTCGGCTCT 605  
DB 466 GCCTGATTCGGGCTTCATGACATTCGATTCCTGTTACTGGTATTCGGCTCT 525

QY 606 GTGACAGTCTACCATATATGTTACAGATATACAGGAAACCGGGTTACTAGTAGCG 665  
DB 526 GTGACAGTCTACCATATATGTTACAGATATACAGGAAACCGGGTTACTAGTAGCG 585

QY 666 CCATAGCTGCAACCTTTCACCTCCACTGCAATGCTGGCCCTGCAAGCTGGGCTGT 725  
DB 586 CCATAGCTGCAACCTTTCACCTCCACTGCAATGCTGGCCCTGCAAGCTGGGCTGT 645

QY 726 TGCCCTGCCCCCTTGGTCCCTGACATACAGCAGTTTATACCCACACACCTGTCTA 785  
DB 646 TGCCCTGCCCCCTTGGTCCCTGACATACAGCAGTTTATACCCACACACCTGTCTA 705

QY 786 CACTGACATTCATTAAGTG-ACGTGCTGTGAAAAAACAATAAAACCC 837  
DB 706 CAGTGTCAATTCAATAAAGTGACGCTGTGTGAAACACAAAAAACAACACAC 758

RESULT 14  
US-10-176-847-27  
; Sequence 27, Application US/10176847  
; Publication No. US2003006836A1  
; GENERAL INFORMATION:  
; APPLICANT: Veiby, Petter Ole  
; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR  
; TITLE OF INVENTION: IDENTIFICATION ASSESSMENT, PREVENTION, AND THERAPY OF BREAST  
; TITLE OF INVENTION: AND OVARIAN CANCER  
; FILE REFERENCE: MRI-039  
; CURRENT APPLICATION NUMBER: US/10/176,847  
; CURRENT FILING DATE: 2002-06-21  
; NUMBER OF SEQ ID NOS: 112

; SOFTWARE: PastSeq for Windows Version 4.0  
; SEQ ID NO 27  
; LENGTH: 852  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-176-847-27

Query Match 61.8%; Score 685.8; DB 14; Length 852;  
Best Local Similarity 99.4%; Pred. No. 3.6e-205; Indels 2; Gaps 2;  
Matches 709; Conservative 0; Mismatches 2;

QY 106 AGCAGGAATAGAAACTTAAGAGAAATACACACTTCTGAGAAACTGAAACGACGGGAA 165  
DB 141 AGCAGGAATAGAAACTTAAGAGAAATACACACTTCTGAGAAACTGAAACGACGGGAA 200

QY 166 AGAGGTCTCACTGAGCAGCGTCCAGCATCCGACACCCAGCGGCGCTTCGCTCCACG 225  
DB 201 AGAGGTCTCACTGAGCAGCGTCCAGCATCCGACACCCAGCGGCGCTTCGCTCCACG 260

QY 226 CAGAAAAACCACTTCTCAAAACCTTCACTCAACACTTCTTCCCAAGCCAGAGATGC 285  
DB 261 CAGAAAAACCACTTCTCAAAACCTTCACTCAACACTTCTTCCCAAGCCAGAGATGC 320

QY 286 ACAGAGAGNACATGAGTGGCTGTGGGGGACCCCGACACCATCTTCCAAAGT 345  
DB 321 ACAAGGAGGAAACATGAGTGGCTGTGGGGGACCCCGACACCATCTTCCAAAGT 380

QY 346 CCAACCGTGATCAACATCCACAGCAGACCTCCGTGCGCCGACCATCTGCTGCTCCCTGT 405  
DB 381 CCAACCGTGATCAACATCCACAGCAGACCTCCGTGCGCCGACCATCTGCTGCTCCCTGT 440

QY 406 TCAACACCTCTTCTTGAACCTGTGCTGGGCTTCATAGCATTCGCTACTCCGCTGA 465  
DB 441 TCAACACCTCTTCTTGAACCTGTGCTGGGCTTCATAGCATTCGCTACTCCGCTGA 500

QY 466 AGTCTAGGACAGGAAGATGTTGGCGACGCTGACCGGGGCCAGGCTATGCTCCACCG 525  
DB 501 AGTCTAGGACAGGAAGATGTTGGCGACGCTGACCGGGGCCAGGCTATGCTCCACCG 560

QY 526 CCAAGTGCCTGACATCTGGGCGCTGATCTGGGCGATCTCATGACCATTCGATTCATCC 585  
DB 561 CCAAGTGCCTGACATCTGGGCGCTGATCTGGGCGATCTCATGACCATTCGATTCATCC 620

QY 586 TGTACTGTGTTATCGGCTCTGTGACAGTCTACCATATATGTTACAGATAATACAGGAAA 645  
DB 621 TGTACTGTGTTATCGGCTCTGTGACAGTCTACCATATATGTTACAGATAATACAGGAAA 680

QY 646 AACGGGTTACTAGTAGCCGCGCATAGCCTGCAACCTTTGCACTCCACTGTGCAATGCTG 705  
DB 681 AACGGGTTACTAGTAGCCGCGCATAGCCTGCAACCTTTGCACTCCACTGTGCAATGCTG 740

QY 706 GCCTTGACGCTGGGGCTGTTGGCCCTGCGCCCTTGGTCTGCTGCTGCTGCTGCTGCTGCT 765  
DB 741 GCCTTGAC-CTGGGGCTGTTGGCCCTGCGCCCTTGGTCTGCTGCTGCTGCTGCTGCTGCT 799

QY 766 TATACCCACACCTGTCTACACTGACATTCATTAAGTG-ACGTGCTGTGTA 817  
DB 800 TATACCCACACCTGTCTACACTGACATTCATTAAGTGACGTGCTGTGTA 852

RESULT 15  
US-10-301-822-80  
; Sequence 80, Application US/10301822  
; Publication No. US20030148410A1  
; GENERAL INFORMATION:  
; APPLICANT: Millennium Pharmaceuticals, Inc.  
; APPLICANT: Berger, Allison  
; APPLICANT: Guillemette, Tracy L.  
; APPLICANT: Kamatkar, Shubhangi  
; APPLICANT: Schlegel, Robert  
; APPLICANT: Monahan, John E.  
; APPLICANT: Thibodeau, Stephen N.  
; APPLICANT: Burgart, Lawrence J.

Job time : 261.529 secs

;; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND  
;; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND  
;; TITLE OF INVENTION: THERAPY OF COLON CANCER  
;; FILE REFERENCE: MP001-0292RNM  
;; CURRENT APPLICATION NUMBER: US/10/301,822  
;; CURRENT FILING DATE: 2002-11-21  
;; PRIOR APPLICATION NUMBER: US 60/339,971  
;; PRIOR FILING DATE: 2001-12-10  
;; PRIOR APPLICATION NUMBER: US 60/361,978  
;; PRIOR FILING DATE: 2002-03-05  
;; PRIOR APPLICATION NUMBER: US 60/381,988  
;; PRIOR FILING DATE: 2002-05-20  
;; NUMBER OF SEQ ID NOS: 228  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 80  
;; LENGTH: 647  
;; TYPE: DNA  
;; ORGANISM: Homo Sapiens  
;; FEATURE:  
;; NAME/KEY: CDS  
;; LOCATION: (111)...(488)  
;; US-10-301-822-80

Query Match 56.3%; Score 624.4; DB 14; Length 647;  
Best Local Similarity 98.9%; Pred. No. 7e-186;  
Matches 639; Conservative 0; Mismatches 6; Indels 1; Gaps 1;  
Qy 173 CTCACCTGAGCAGCGTCCGAGCATCCGAGACACACAGCGGCGCTTCGCTCCACGCGAGAAA 232  
Db 2 CTCACCTGAGCAGCGTCCGAGCATCCGAGACACACAGCGGCGCTTCGCTCCACGCGAGAAA 61  
Qy 233 CCACACTTCTCAAACTTCACCTCACTCAACACTTCTTCCTCCCAAAGCCAGAGATGCACAAAGGA 292  
Db 62 CCACACTTCTCACTTCTCACTCAACACTTCTTCCTCCCAAAGCCAGAGATGCACAAAGGA 121  
Qy 293 GGAACATGAGGTGGTGTGCTGGGGGACACCCCGAGCACCATTCTTCCTCCAGGTCCACCGT 352  
Db 122 GGAACATGAGGTGGTGTGCTGGGGGACACCCCGAGCACCATTCTTCCTCCAGGTCCACCGT 181  
Qy 353 GATCAACATCCACAGGAGACCTCCGTCGCCGACCATGTCGTGTGGTCCCTGTTCAACAC 412  
Db 182 GATTAACATCCACAGGAGACCTCCGTCGCCGACCATGTCGTGTGGTCCCTGTTCAACAC 241  
Qy 413 CCTCTTCTTGAAGTGTGCTGTGGGCTTCATAGCATTCGCTACTCCGTTGAGTCTAG 472  
Db 242 CCTCTTCTTGAAGTGTGCTGTGGGCTTCATAGCATTCGCTACTCCGTTGAGTCTAG 301  
Qy 473 GGACAGGAAGATGGTTGGGACGTCGACCGGGGCCAGGCGCTATGCCCTCCACCGCCAAAGTG 532  
Db 302 GGACAGGAAGATGGTTGGGACGTCGACCGGGGCCAGGCGCTATGCCCTCCACCGCCAAAGTG 361  
Qy 533 CCTGAACATCTGGGCGCTGATTCCTGGGCGCTCCTCATGACATTGGATTATCCTGTTACT 592  
Db 362 CCTGAACATCTGGGCGCTGATTCCTGGGCGCTCCTCATGACATTGGATTATCCTGTTACT 421  
Qy 593 GGTATTGGGCTCTGTACAGTCTACCATATTATGTTACAGATATACAGGAAAACGGGG 652  
Db 422 GGTATTGGGCTCTGTACAGTCTACCATATTATGTTACAGATATACAGGAAAACGGGG 481  
Qy 653 TTACTAGTAGCCGCCCATAGCCTGCAACCTTTGCACTTCCACTGCAATGTGCGCCCTGC 712  
Db 482 TTACTAGTAGCCGCCCATAGCCTGCAACCTTTGCACTTCCACTGCAATGTGCGCCCTGC 541  
Qy 713 ACGCTGGGCTGTGGCCCTGCGCCCTTGGTCCCTGCTAGATACAGCAGTTTATACCC 772  
Db 542 ACGCTGGGCTGTGGCCCTGCGCCCTTGGTCCCTGCTAGATACAGCAGTTTATACCC 601  
Qy 773 ACACACCTGTCTACACTGACATTCATTAAGTG-ACGTGTTGTGA 817  
Db 602 ACACACCTGTCTACAGTGTCTTCAATAAAGTGACGCTGTGTGA 647

Search completed: March 12, 2004; 06:15:10



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 10, 2004, 09:18:38 ; Search time 1836.07 Seconds

(without alignments)  
18036.975 Million cell updates/sec

Title: US-10-084-817-37

Perfect score: 1109

Sequence: 1 cctgaccaggagacactgg.....acanaattccacctatccg 1109

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: em\_estha:\*

2: em\_esthm:\*

3: em\_estin:\*

4: em\_estnu:\*

5: em\_estov:\*

6: em\_estpl:\*

7: em\_estro:\*

8: em\_htc:\*

9: gb\_est1:\*

10: gb\_est2:\*

11: gb\_htc:\*

12: gb\_est3:\*

13: gb\_est4:\*

14: gb\_est5:\*

15: em\_estfun:\*

16: em\_estom:\*

17: em\_gss\_hum:\*

18: em\_gss\_inv:\*

19: em\_gss\_pln:\*

20: em\_gss\_vit:\*

21: em\_gss\_fun:\*

22: em\_gss\_mam:\*

23: em\_gss\_mus:\*

24: em\_gss\_pro:\*

25: em\_gss\_fod:\*

26: em\_gss\_phg:\*

27: em\_gss\_vrl:\*

28: gb\_gss1:\*

29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	726.8	65.5	841	12	BI909898 603070530
2	689.8	62.2	727	14	CD174424 AGENCOURT
3	685	61.8	835	13	BU603243 AGENCOURT
4	682.8	61.6	728	14	CD174239 AGENCOURT

5	682.2	61.5	718	14	CF780624
6	678.2	61.2	717	14	CD175008
7	678	61.1	726	14	CD175837
8	676.2	61.0	718	14	CD172759
9	672.8	60.7	726	14	CD176014
10	671.8	60.6	721	14	CD175245
11	668.8	60.3	727	9	AV759074
12	668	60.2	704	14	CD173064
13	667.2	60.2	715	14	CD173397
14	666.6	60.1	713	14	CD173073
15	666.4	60.1	712	14	CD175286
16	665.4	60.0	710	14	CD172976
17	665	60.0	710	14	CD176886
18	665	60.0	710	14	CD247128
19	664.6	59.9	691	13	BQ219055
20	663	59.8	899	14	CD250223
21	662	59.7	699	12	BG719484
22	661.2	59.6	709	14	CD175451
23	661	59.6	715	14	CD172777
24	661	59.6	726	14	CD249234
25	660.6	59.6	713	14	CD174839
26	660	59.5	711	14	CD247389
27	659	59.4	738	13	BU537783
28	658.2	59.4	707	14	CD173531
29	657.6	59.3	691	13	BU558290
30	655.8	59.1	691	13	BU558290
31	653.4	58.9	700	14	CD247008
32	653	58.9	707	14	CD173309
33	652.6	58.8	793	9	AV758898
34	651.8	58.8	703	14	CD175270
35	651.6	58.8	670	12	BM995109
36	651	58.7	699	14	CD385219
37	650.2	58.6	669	12	BG534713
38	649.6	58.6	697	14	CD702719
39	649.4	58.6	704	14	CD248465
40	649	58.5	712	14	CD512187
41	647	58.3	895	13	BQ228236
42	646.8	58.3	762	10	BF665518
43	641.2	57.8	662	9	AV759976
44	639.6	57.7	706	14	CA429526
45	639.2	57.6	673	14	CB151066

ALIGNMENTS

RESULT 1  
BI909898  
LOCUS  
DEFINITION 603070530F1 NIH\_MGC\_118 Homo sapiens CDNA clone IMAGE:5219479 5',  
mRNA sequence.  
ACCESSION BI909898  
VERSION B:909898.1 GI:16173196  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 841)  
NIH-MGC http://mgc.ncbi.nlm.nih.gov/.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
AUTHORS Contact: Robert Strausberg, Ph.D.  
TITLE Email: csapbs-remail.nih.gov  
JOURNAL Tissue Procurement: Life Technologies, Inc.  
COMMENT CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LINL at:  
http://image.llnl.gov  
Plate: LHAM11551 row: 1 column: 08  
High quality sequence stop: 824.

```
FEATURES
source
Location/Qualifiers
1. 841
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5219479"
/tissue_type="leukocyte"
/lab_host="DH10B"
/clone_lib="NIH_MGC 118"
/note="Vector: pCMV-SORT6; Site 1: NotI; Site 2: EcoRV
(deconstructed); RNA source leukocytes from anonymous pool of
non-activated adult donors. Library is oligo-dT primed
and directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.7 kb, insert size range
1.2-3.3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(invitrogen). Research Genetics tracking code 027. Note:
this is a NIH_MGC Library."

ORIGIN
Query Match 65.5%; Score 726.8; DB 12; Length 841;
Best Local Similarity 99.3%; Pred. No. 5.6e-137;
Matches 761; Conservative 0; Mismatches 2; Indels 3; Gaps 3;

QY 1 CCTGCACGAGGACACTGGAGGTTAGTCCCAACCCGACAGAGGAGGACTGCAGC 60
Db 59 CCTGCACGAGGACACTGGAGGTTAGTCCCAACCCGACAGAGGAGGACTGCAGC 118
QY 61 CTGAGGAAGAGCAAGATTTACGAGAGAGGCTCGGACAAAGTGACAGGAATAGAAA 120
Db 119 CTGAGGAAGAGCAAGATTTACGAGAGAGGCTCGGACAAAGTGACAGGAATAGAAA 178
QY 121 CTTAAGAGAAATACACACTTCTCAGAAACTGAAACGACAGGGGAAAGGAGTCTCACTGA 180
Db 179 CTTAAGAGAAATACACACTTCTCAGAAACTGAAACGACAGGGGAAAGGAGTCTCACTGA 238
QY 181 GCACGCTCCAGCATCCGACACACACAGCGGCGCTTCGCTCCAGCGAGAAACACACTT 240
Db 239 GCACGCTCCAGCATCCGACACACACAGCGGCGCTTCGCTCCAGCGAGAAACACACTT 298
QY 241 CTCACACCTTCACTCAACACTTCTTCCCAAGCCAGAGATGCAACAGGAGAAACATG 300
Db 299 CTCACACCTTCACTCAACACTTCTTCCCAAGCCAGAGATGCAACAGGAGAAACATG 358
QY 301 AGGTGGCTGTGTGGGGGACCCCGACAGACCACTTCCTTCAAGTCCACCGTGATCAACA 360
Db 359 AGGTGGCTGTGTGGGGGACCCCGACAGCACCCTCTTCAAGTCCACCGTGATCAACA 418
QY 361 TCCACAGCAGACCTCGTCCCGGACCATGTCGTCCTGTCCTGTCCTTCAACACCTTCT 420
Db 419 TCCACAGCAGACCTCGTCCCGGACCATGTCGTCCTGTCCTGTCCTTCAACACCTTCT 478
QY 421 TGAACCTGGTGTCTGTGGGCTTCATAGCAATTCGCTTACCTCGTGAAGTCTAGGGACGGA 480
Db 479 TGAACCTGGTGTCTGTGGGCTTCATAGCAATTCGCTTACCTCGTGAAGTCTAGGGACGGA 538
QY 481 AGATGTTGGCGAGCTGACCGGGGCCCCGACCTATGCTCCACCGCAAGTGCCTGGAACA 540
Db 539 AGATGTTGGCGAGCTGACCGGGGCCCCGACGCTATGCTCCACCGCAAGTGCCTGGAACA 598
QY 541 TCTGGGCCCCGTATCTGGGGCATCTCATGACCAATTGGATTTCATCTGTACTGGTATTTCG 600
Db 599 TCTGGGCCCCGTATCTGGGGCATCTCATGACCAATTGGATTTCATCTGTACTGGTATTTCG 658
QY 601 GCTCTGTGACAGTCTACCATATTATGTACAGATAATACAGGAAACGGGGTACTAGT 660
Db 659 GCTCTGTGACAGTCTACCATATTATGTACAGATAATACAGGAAACGGGGTACTAGT 718
QY 661 AGCGGCCCATAGCTGCAACTTTCGACTCCACTG-TGCAATGCTGGCCCTGCACGCTGG 719
Db 719 AGCGG-CCATAGCTGCAACTTTCGACTCCACTGTTGCAATGCTGGCCCTGCACGCTGG 777
QY 720 GGCTGTGCCCCCTTGGTCTTGGCCCCCTAGATACAGCAGTT 765
```

```
Db 778 GGCTGTGCCCCCTG-CCCTTGGTCTTGGCCCCCTAGATACAGCAGTT 822

RESULT 2
CD174424
LOCUS CD174424 727 bp mRNA linear EST 19-MAY-2003
DEFINITION AGENCOURT_13973296 NIH_MGC_172 Homo sapiens cDNA 5', mRNA sequence.
ACCESSION CD174424
VERSION CD174424.1 GI:30857965
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 727)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: Dr. Jamie Thompson, University of WI
cDNA Library Preparation: Gina Zastrow-Hayes
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: NDRM48 row: p column: 16
High quality sequence start: 9
High quality sequence stop: 727.

FEATURES
source
1. 727
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="embryonic stem cells, WA01, passage 38"
/lab_host="DH10B Tona"
/clone_lib="NIH_MGC 172"
/note="Vector: pDONR201; Site 1: attP2; Site 2: attP1;
L1BR PRIMING - oligo HT; METHOD - full-length enriched;
Embryonic Stem Cells H1; L1BR PROVIDER - Bradfield"

ORIGIN
Query Match 62.2%; Score 689.8; DB 14; Length 727;
Best Local Similarity 98.1%; Pred. No. 1.7e-139; Indels 1; Gaps 1;
Matches 708; Conservative 0; Mismatches 13;

QY 117 GAAACTTAAGAGAAATACACACTTCTGAGAAACTGAAACGACAGGGGAAAGAGGTCTCA 176
Db 5 GAAGGCTAGAAANATACACACTTCTGAGAAACTGAAACGACAGGGGAAAGAGGTCTCA 64
QY 177 CTGAGCACCGTCCAGCATCCGACACACAGCGGCGCTTCGCTCCACGAGAAACAC 236
Db 65 CTGAGCACCGTCCAGCATCCGACACACAGCGGCGCTTCGCTCCACGAGAAACAC 124
QY 237 ACTTCTCAAACTTCACTCAACACTTCTTCTTCCCAAGCCAGAGATGCAACAGGAGAA 296
Db 125 ACTTCTCAAACTTCACTCAACACTTCTTCTTCCCAAGCCAGAGATGCAACAGGAGAA 184
QY 297 CATGAGGTGGCTGTGTGGGGGCACTCCGTCGCGGACCATCTTCCAGGTCCACCGTGATC 356
Db 185 CATGAGGTGGCTGTGTGGGGGCACTCCGTCGCGGACCATCTTCCAGGTCCACCGTGATC 244
QY 357 AACATCCACAGCAGACCTCCGTCGCGGACCATCTTCCAGGTCCACCGTGATC 416
Db 245 AACATCCACAGCAGACCTCCGTCGCGGACCATCTTCCAGGTCCACCGTGATC 304
QY 417 TTCTTGAACHTGGTGTCTGTGGGGTTCATAGCATTCGCTACTCGTGAAGTCTAGGGAC 476
Db 305 TTCTTGAACHTGGTGTCTGTGGGGTTCATAGCATTCGCTACTCGTGAAGTCTAGGGAC 364
QY 477 AGGAAGATGTTGGCGACGTGACCGGGGCCAGGCGCTATGCTCCACCGCAAGTGCCTG 536
```

```

365 AGAAGATGGTGTGGCAGCTGACCGGGGCCAGGGCTATGCTCCACCGCCAAAGTGCCTG 424
537 AACATCTGGGCGCTGATCTGGGATCCTCATGACATTTGATTCATCTCTGTTACTGGTA 596
425 AACATCTGGGCGCTGATCTGGGATCCTCATGACATTTGATTCATCTCTGTTACTGGTA 484
597 TTGGGCTCTGTGACAGTCTACATATTTATGTTACAGATAAATACAGAAACCGGGTTAC 656
485 TTGGGCTCTGTGACAGTCTACATATTTATGTTACAGATAAATACAGAAACCGGGTTAC 544
657 TAGTAGCGCGCCATAGCTGCAACTTTCACCTCCACTGTCGCAATGCTGGCCCTGCACGC 716
545 TAGTAGCGCGCCATAGCTGCAACTTTCACCTCCACTGTCGCAATGCTGGCCCTGCACGC 604
717 TGGGGCTGTGGCCCTGCCCTTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 776
605 TGGGGCTGTGGCCCTGCCCTTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 664
777 ACCTGTCTACACTGACATTCATTAAGTG-AGTGCTTGTGAAAGAAACAAATTAAC 835
665 ACCTGTCTACAGTGTCAATTAAGTGACGCTGCTTGTGAAAGAAACAAATTAAC 724
836 CC 837
725 CC 726

```

RESULT 3  
 BU603243  
 LOCUS  
 DEFINITION AGENCOURT 10059831 NIH MGC 142 Homo sapiens cdna clone  
 IMAGE:6498263 5', mRNA sequence.

ACCESSION BU603243  
 VERSION BU603243.1 GI:23255002  
 KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 835)

NIH-MGC <http://mgi.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)

Tissue Procurement: NCI

CDNA Library Preparation: Michael Brownstein Laboratory

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LICM2680 row: j column: 24

High quality sequence stop: 558.

Location/Qualifiers

1. 835

FEATURES

source

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:6498263"

/tissue\_type="mixed (pool of 40 RNAs)"

/lab\_hosts="DH10B (TI-phage-resistant)"

/clone\_lib="NIH\_MGC\_142"

/note="Vector: pDNR-LIB; Site 1: Sfil (ggcattatggcc);

Site 2: Sfil (ggcgctctggcc); Double-stranded cDNA was

prepared from a pool of 40 cell line polyA+ RNAs (bladder

- 2%, blood - 33.4%, brain - 5.6%, breast - 12.5%, colon -

4%, connective tissue - 1.4%, eye - 1%, intestine - 2.6%,

kidney - 2.2%, liver - 5.7%, lung - 10.8%, NK-cell - 5.2%,

ovary - 4%, pharynx - 2.5%, prostate - 4.3%, salivary

gland - 1.3%, and skin - 2.3%). 5' and 3' adaptors were

used in cloning as follows:

5'-AAGCAGTGGTATCAACGACAGAGTGGCCATTACGCCGGG-3' and  
 5'-ATTCTAGAGCGGAGCGGCCGACATG-3' (30)NN-3'. Full-length  
 enriched library was constructed using the Clontech  
 Creator SMART kit and size-selected to contain the >0.5 kb  
 size fraction (other fractions present in NIH MGC\_141).  
 Library created in the laboratory of M. Brownstein (NIMH,  
 NIH). Note: this is a NIH\_MGC Library."

ORIGIN

Query Match 61.8%; Score 685; DB 13; Length 835;  
 Best Local Similarity 96.6%; Pred. No. 1.6e-128;  
 Matches 705; Conservative 0; Mismatches 24; Indels 1; Gaps 1;

QY	106	AGCAGGAAATAGAACTTAAAGAGAAATACACTTCTGAGAACTGAAACGACAGGGGAA	165
DB	8	AGCAGGAAATAGAACTTAAAGAGAAATACACTTCTGAGAACTGAAACGACAGGGGAA	67
QY	166	AGGAGGTCTCACTGAGCACCCTCCAGCATCCGACACACACAGCGGCCCTTCCTCCACG	225
DB	68	AGGAGGTCTCACTGAGCACCCTCCAGCATCCGACACACACAGCGGCCCTTCCTCCACG	127
QY	226	CAGAAAACCACTTCTCAAACTTCACTCAACACTTCTTCCCAAGCAGAAATGC	285
DB	128	CAGAAAACCACTTCTCAAACTTCACTCAACACTTCTTCCCAAGCAGAAATGC	187
QY	286	ACAAGGAGGACATGAGTGGCTGTGGGGGACCCCGACGACCATCTTCCAGGT	345
DB	188	ACAAGGAGGACATGAGTGGCTGTGGGGGACCCCGACGACCATCTTCCAGGT	247
QY	346	CCACCGTGATCAACATCCACAGCAGACCTCCGTGCCCGACCATGTCGTGTGCTCCTGT	405
DB	248	CCACCGTGATCAACATCCACAGCAGACCTCCGTGCCCGACCATGTCGTGTGCTCCTGT	307
QY	406	TCAACACCTCTTCTTGAACCTGCTGTGGGCTTCACTAGCATTCGCTACTCCGTGA	465
DB	308	TCAACACCTCTTCTTGAACCTGCTGTGGGCTTCACTAGCATTCGCTACTCCGTGA	367
QY	466	AGTCTAGGGACAGGAAGATGGTGGCGACGTGACCGGGGCCAGGCGCTATGCCTCCACCG	525
DB	368	AGTCTAGGGACAGGAAGATGGTGGCGACGTGACCGGGGCCAGGCGCTATGCCTCCACCG	427
QY	526	CCAAAGTCCCTGAACATCTGGGCCCTGATCTGGGCATCTCTCATGACCATTCATTC	585
DB	428	CCAAAGTCCCTGAACATCTGGGCCCTGATCTGGGCATCTCTCATGACCATTCATTC	487
QY	586	TGTTACTGGTATTCGGCTCTGTGACAGTCTACCATATTATGTACAGATAATACA-GGAA	644
DB	488	TGTTACTGGTATTCGGCTCTGTGACAGTCTACCATATTATGTACAGATAATACA-GGAA	547
QY	645	AAACGGGGTTACTAGTAGCCGCCATAGCCTGCAACCTTTGCACTCCACTGTGCAATGCT	704
DB	548	AAACGGGGTTACTAGTAGCCGCCATAGCCTGCAACCTTTGCACTCCACTGTGCAATGCT	607
QY	705	GGCCCTGCACGCTGGGGCTGTGGCCCTGCCCTTGGTCTGCTGCCCTAGATACAGAGT	764
DB	608	GGCCCTGCACGCTGGGGCTGTGGCCCTGCCCTTGGTCTGCTGCCCTAGATACAGAGT	667
QY	765	TTATACCCACACACCTGTCTACACTGACATTCATTAAGTGCAGTGTGTGAAAAAAA	824
DB	668	TTATACCCACACACCTGTCTACAGTGTCTATTTCATAAAGTGCAGCGTGTGTGANNAA	727
QY	825	ACAAATAAAA 834	
DB	728	AANNANNA 737	

RESULT 4  
 CD174239  
 LOCUS  
 DEFINITION AGENCOURT\_13976705 NIH\_MGC\_172 Homo sapiens CDNA 5', mRNA sequence.  
 ACCSSION CD174239  
 VERSION CD174239.1 GI:30857594  
 KEYWORDS EST.



117	QA	AACTTTAAGAGAAATACACACTTCTTGAGAAACTGAAACGACAGGGGAAAGAGAGTCTCA	176
7	DB	GAACCTTTAAGAGAAATACACACTTCTTGAGAAACTGAAACGACAGGGGAAAGAGAGTCTCA	66
177	QY	CTGAGCACCGTCCCAGCATCCGACACACACAGCGGCCCTTCGCTCCACGACAGAAACACAC	236
67	DB	CTGAGCACCGTCCCAGCATCCGAGACACACAGCGGCCCTTCGCTCCAACGAGAAACACAC	126
237	QY	ACTTCTCAAAACCTTTCATCAACACTTCTTCCCAAGCCAGAGATGCACAGAGAGAA	296
127	DB	ACTTCTCAAAACCTTTCATCAACACTTCTTCCCAAGCCAGAGATGCACAGAGAGAA	186
297	QY	CATGAGGTGGCTGTCTGGGGCACCCGCCAGCACCATCTCTTCCAAGTTCACCGTGTATC	356
187	DB	CATGAGGTGGCTGTCTGGGGCACCCGCCAGCACCATCTCTTCCAAGTTCACCGTGTATC	246
357	QY	AACATCCACAGCGAGACCTCCGTGCCCGACCATGTCTGCTCCCTGCTTCAACACCCCTC	416
247	DB	AACATCCACAGCGAGACCTCCGTGCCCGACCATGTCTGCTCCCTGCTTCAACACCCCTC	306
417	QY	TTCTTGAACCTGGTGCTGTCTGGGCTTCACTAGCATTTCCGCTACTCCGTCGAAGTCTAGGAC	476
307	DB	TTCTTGAACCTGGTGCTGTCTGGGCTTCACTAGCATTTCCGCTACTCCGTCGAAGTCTAGGAC	366
477	QY	AGGAAGATGGTTGGCGAGCTGACCGGGGCCAGGSCCTATGCTCCACGCCCAAGTGCCTG	536
367	DB	AGGAAGATGGTTGGCGAGCTGACCGGGGCCAGGSCCTATGCTCCACGCCCAAGTGCCTG	426
537	QY	AACATCTGGGCCCTGATTCTGGGCATCTCATGACCATTTGGATTCACTCTGTACTGTGTA	596
427	DB	AACATCTGGGCCCTGATTCTGGGCATCTCATGACCATTTGGATTCACTCTGTACTGTGTA	486
597	QY	TTCCGCTCTGTGACAGTCTACCATATTATTTACAGATAATACAGAAAAACGGGGTTAC	656
487	DB	TTCCGCTCTGTGACAGTCTACCATATTATTTACAGATAATACAGAAAAACGGGGTTAC	546
657	QY	TAGTAGCGGCCCATAGCCTGAACTTTTGCACTCCACTGTGCAATGCTGGCCCTGCACGC	716
547	DB	TAGTAGGCCCTCCATAGCCTGAACTTTTGCACTCCACTGTGCAATGCTGGCCCTGCACGC	606
717	QY	TGGGGCTGTTGCCCTGCCCCCTTGGTCTGCCCCCTAGATACAGCAGTTTATACCCACAC	776
607	DB	TGGGGCTGTTGCCCTGCCCCCTTGGTCTGCCCCCTAGATACAGCAGTTTATACCCACAC	666
777	QY	A-CTGTCTACACTGACATTCAAATAAAGTG-ACGTGCTTTGTAAAAAATAA	825
667	DB	ACCGTGTCTACAGTGTCAATCAATAAGTGACGTGCTTTGTAAAAAATAA	717

RESULT 6  
CDI175008  
LOCUS  
CDI175008  
DEFINITION  
ACCESSION  
AGNCOURT\_13971378 NIH\_MGC\_172 Homo sapiens cDNA 5', mRNA sequence.  
VERSION  
CDI175008.1 GI:30859148  
KEYWORDS  
EST.  
SOURCE  
CDI175008  
Homo sapiens (human)  
ORGANISM  
Homo sapiens  
Eukaryota; Chordata; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
REFERENCE  
1 (bases 1 to 717)  
NHL-MGC <http://mgc.nci.nih.gov/>.  
AUTHORS  
National Institutes of Health, Mammalian Gene Collection (MGC)  
TITLE  
Unpublished (1999)  
JOURNAL  
Contact: Robert Strausberg, Ph.D.  
COMMENT  
Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)  
Tissue Procurement: Dr. Jamie Thompson, University of WI  
cDNA Library Preparation: Gina Zastrow-Hayes  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>

Plate: NDKM29 row: c column: 01  
 High quality sequence stop: 717.  
 Location/Qualifiers  
 1. .717  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /tissue\_type="embryonic stem cells, WA01, passage 38"  
 /lab\_host="DH10B TONa"  
 /clone\_lib="NIH MGC 172"  
 /note="Vector: pDONR201; Site 1: attP2; Site 2: attP1;  
 LIBR PRIMING - Oligo 4T; METHOD - full-length enriched;  
 Embryonic Stem Cells H1; LIBR PROVIDER - Bradfield"

ORIGIN

Query Match 61.2%; Score 678.2; DB 14; Length 717;  
 Best Local Similarity 98.0%; Pred. No. 3.9e-127;  
 Matches 697; Conservative 0; Mismatches 13; Indels 1; Gaps 1;

QY 128 GAAATACACACTTCTGAGAAACTGAAACGACAGGGGAAAGAGGTCTCACTGAGCACCGT 187  
 |||||  
 Db 6 GAAGGAGTAGCACTGAGAACTGAAACGACAGGGGAAAGAGGTCTCACTGAGCACCGT 65  
 |||||

QY 188 CCCAGCATCGGACACACAGCGGCCCTTCGCTCCAGCAAGAAAACACACTTCTTCAAAC 247  
 |||||

Db 66 CCCAGCATCGGACACACAGCGGCCCTTCGCTCCAGCAAGAAAACACACTTCTTCAAAC 125  
 |||||

QY 248 CTTCACTCAACACTTCTCTCCCAAGCCAGAGATGCACAAGAGAAACATCAGGTGGC 307  
 |||||

Db 126 CTTCACTCAACACTTCTCTCCCAAGCCAGAGATGCACAAGAGAAACATCAGGTGGC 185  
 |||||

QY 308 TGTGCTGGGGGACACCCCGCAGCACCATTCCTTCCAAGTCCACCGTGATCAACATCCAAG 367  
 |||||

Db 186 TGTGCTGGGGGACACCCCGCAGCACCATTCCTTCCAAGTCCACCGTGATCAACATCCAAG 245  
 |||||

QY 368 CGAGACTCCGTCGCGGACCATGCTGCTCGCTGCTTCAACACCCCTCTCTTGAACTG 427  
 |||||

Db 246 CGAGACTCCGTCGCGGACCATGCTGCTGCTGCTTCAACACCCCTCTCTTGAACTG 305  
 |||||

QY 428 GTGCTGTCTGGGCTTCATAGCATTCGCTACTCCGTGAACTCTAGGACAGGAAGATGGT 487  
 |||||

Db 306 GTGCTGTCTGGGCTTCATAGCATTCGCTACTCCGTGAACTCTAGGACAGGAAGATGGT 365  
 |||||

QY 488 TGGCGAGGTGACCGGGGCCAGGCGCTATGCTCCACCGCCAGTGCCTGACATCTGGC 547  
 |||||

Db 366 TGGCGAGGTGACCGGGGCCAGGCGCTATGCTCCACCGCCAGTGCCTGACATCTGGC 425  
 |||||

QY 548 CCTGATTCTGGGCATCTCTCATGACCATTTGATTATCATCTGTACTGTATTCTGGCTCTGT 607  
 |||||

Db 426 CCTGATTCTGGGCATCTCTCATGACCATTTGATTATCATCTGTACTGTATTCTGGCTCTGT 485  
 |||||

QY 608 GACAGTCAACATATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCGGCC 667  
 |||||

Db 486 GACAGTCAACATATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCGGCC 545  
 |||||

QY 668 CATAGCTGCAACTTTGCATCCACTGTGCAATGCTGGCCCTGACGCTGGGCGCTGTG 727  
 |||||

Db 546 CATAGCTGCAACTTTGCATCCACTGTGCAATGCTGGCCCTGACGCTGGGCGCTGTG 605  
 |||||

QY 728 CCCTGCGCCCTTGGTCTCCCTAGATACAGCAGTTTATACCCACACCTGCTTACA 787  
 |||||

Db 606 CCCTGCGCCCTTGGTCTCCCTAGATACAGCAGTTTATACCCACACCTGCTTACA 665  
 |||||

QY 788 CTGACATTCATTAAGTG-ACGTGCTGTGAAAAAACAACATAAAACCC 837  
 |||||

Db 666 GTGTCAATCAATGAAGTGCAAGTCTGTGAAAAAACAACATAAAACCC 716  
 |||||

RESULT 7  
 CD175837  
 LOCUS CD175837 726 bp mRNA linear EST 19-MAY-2003  
 DEFINITION AGENCOURT\_13976863 NIH\_MGC\_172 Homo sapiens cDNA 5', mRNA sequence  
 ACCESSION CD175837

CD175837.1	GI:30860834	Db	484	GGCTCTGTGCACAGTCTACCATATTATGTTACAGATATATACAGGAAAAACGGGGTTACTAG	543
EST.					
KEYWORDS					
ORGANISM					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
COMMENT					
NIH-MGC	http://mgs.nci.nih.gov/				
1	(bases 1 to 726)				
National Institutes of Health, Mammalian Gene Collection (MGC)					
Unpublished (1999)					
Contact: Robert Strausberg, Ph.D.					
Email: cgabbs-r@mail.nih.gov					
Tissue Procurement: Dr. Jamie Thompson, University of WI					
CDNA Library Arrayed by: Gena Zastrow-Hayes					
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)					
DNA Sequencing by: Agencourt Bioscience Corporation					
Clone distribution: MGC clone distribution information can be					
found through the I.M.A.G.E. Consortium/LLNL at:					
http://image.llnl.gov					
Plate: NDKM51 row: m column: 03					
High quality sequence start: 16					
High quality sequence stop: 718.					
Location/Qualifiers					
1. .726					
/organism="Homo sapiens"					
/mol_type="mRNA"					
/db_xref="taxon:9606"					
/tissue_type="embryonic stem cells, WA01, passage 38"					
/lab_host="DH10B Tona"					
/clone_lib="NIH_MGC_172"					
/notes="Vector: pDONR201; Site 1: attP2; Site 2: attP1;					
LIBR PRIMING - oligo dt; METHOD - full-length enriched;					
Embryonic Stem Cells H1; LIBR PROVIDER - Bradfield"					
ORIGIN					
Query Match	61.1%;	Score	678;	DB 14;	Length 726;
Best Local Similarity	97.5%;	Pred. No.	4.3e-127;		
Matches	698;	Conservative	0;	Mismatches	17;
Indels	1;	Gaps	1;		
QY	120	ACTTATAGAGAAATACACACTCTGTGAGAACTGAAACGACAGGGGAAAGGAGGTCTCACTG	179		
DB	4	ACCGAAGAGTAGAAAACCTTCTGAGAACTGAAACGACAGGGGAAAGGAGGTCTCACTG	63		
QY	180	AGCACCGTCCAGCATCCGGACACACACAGCGGCCCTTCGCTCCACGAGAAAAACCACT	239		
DB	64	AGCACCGTCCAGCATCCGGACACACACAGCGGCCCTTCGCTCCACGAGAAAAACCACT	123		
QY	240	TCTCAAACTTCTACTCAACATTCCTTCCCAAGACGACAGAGATGACACAGGAGGACAT	299		
DB	124	TCTCAAACTTCTACTCAACATTCCTTCCCAAGACGACAGAGATGACACAGGAGGACAT	183		
QY	300	GAGTGTGCTGTGTGGGGGACACCCCGACACCATCTTCCAAAGGTCCACCGTGATCAAC	359		
DB	184	GAGTGTGCTGTGTGGGGGACACCCCGACACCATCTTCCAAAGGTCCACCGTGATCAAC	243		
QY	360	ATCCACAGCAGACCTCCGTGCGCGACCATGTGCTGTGTCCTTCCTTCAACACCTCTTC	419		
DB	244	ATCCACAGCAGACCTCCGTGCGCGACCATGTGCTGTGTCCTTCCTTCAACACCTCTTC	303		
QY	420	TTGAACCTGGTGTCTGCGGCTTCATAGCATTCGCTTACCTCCGTTAGGTCTTAGGACAGG	479		
DB	304	TTGAACCTGGTGTCTGCGGCTTCATAGCATTCGCTTACCTCCGTTAGGTCTTAGGACAGG	363		
QY	480	AAGATGTTGGCGAGTGAACCGGGGCCAGGCGCTATGCCCTCCACCGCAAGTGCCTGAAC	539		
DB	364	AAGATGTTGGCGAGTGAACCGGGGCCAGGCGCTATGCCCTCCACCGCAAGTGCCTGAAC	423		
QY	540	ATCTGGGCGCTGATCTTGGGCACTCTCATGACCATTCGATTCATCTCTGTTACTGGTATTC	599		
DB	424	ATCTGGGCGCTGATCTTGGGCACTCTCATGACCATTCGATTCATCTCTGTTACTGGTATTC	483		
QY	600	GGCTCTGTGACAGTCTTACCATTATTTATGATACAGATAATACAGGAAAAACCGGGTTACTAG	659		
CD172759	NIH_MGC_13966328	718 bp	mrna	linear	EST 19-MAY-2003
AGENCOURT					
CD172759	1	GI:30854640			
EST.					
Homo sapiens (human)					
Homo sapiens					
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
1	(bases 1 to 718)				
NIH-MGC	http://mgs.nci.nih.gov/				
National Institutes of Health, Mammalian Gene Collection (MGC)					
Unpublished (1999)					
Contact: Robert Strausberg, Ph.D.					
Email: cgabbs-r@mail.nih.gov					
Tissue Procurement: Dr. Jamie Thompson, University of WI					
CDNA Library Arrayed by: Gena Zastrow-Hayes					
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)					
DNA Sequencing by: Agencourt Bioscience Corporation					
Clone distribution: MGC clone distribution information can be					
found through the I.M.A.G.E. Consortium/LLNL at:					
http://image.llnl.gov					
Plate: NDKM28 row: d column: 04					
High quality sequence start: 11					
High quality sequence stop: 713.					
Location/Qualifiers					
1. .718					
/organism="Homo sapiens"					
/mol_type="mRNA"					
/db_xref="taxon:9606"					
/tissue_type="embryonic stem cells, WA01, passage 38"					
/lab_host="DH10B Tona"					
/clone_lib="NIH_MGC_172"					
/notes="Vector: pDONR201;					



195	TGCTGTCTGGGGGCA	CCCCCCAGCACCATCTCTTCCAAAGGTCCACCGTGATCAACATCCA	254
365	CAGCGAGACCTCCGTGCGCCGACCA	TGTCGTCTGGTCCCTGTTCACACCCCTCTTCTTGAA	424
255	CAGCGAGACCTCCGTGCGCCGACCA	TGTCGTCTGGTCCCTGTTCACACCCCTCTTCTTGAA	314
425	CTGGTGTCTGTCTGGGCTTCA	TAGCATTGCGCTACTCCGTGGAAGTCTAGGGACAGGAAGAT	484
315	CTGGTGTCTGTCTGGGCTTCA	TAGCATTGCGCTACTCCGTGGAAGTCTAGGGACAGGAAGAT	374
485	GGTTGGCGGACGTGACCGGGGGCC	CAGGCGCTATGCCTCCACCGCCCAAGTGCCTGAACATCTG	544
375	GGTTGGCGGACGTGACCGGGGGCC	CAGGCGCTATGCCTCCACCGCCCAAGTGCCTGAACATCTG	434
545	GGCCCTGATTCTGGGCATCCTCAT	GCAATTCGATTCCTGTGTACTGTTATTCGGCTC	604
435	GGCCCTGATTCTGGGCATCCTCAT	GCAATTCGATTCCTGTGTACTGTTATTCGGCTC	494
605	TGTGACAGTCTACCATATATATGTT	TACAGATAATACAGAAAAAAGGGGTACTAGTAGCC	664
495	TGTGACAGTCTACCATATATATGTT	TACAGATAATACAGAAAAAAGGGGTACTAGTAGCC	554
665	GCCCATAGCTGGAACCTTTGCAC	TCCACTGCGAATGCTGGCCCTGCACGCTGGGGGTG	724
555	GCCCATAGCTGGAACCTTTGCAC	TCCACTGCGAATGCTGGCCCTGCACGCTGGGGGTG	614
725	TTGGCCCTGCCCCCTTGCTCTGCT	GCCCTTAGATACAGCAGTTTATATACCAACACACCTGTGCT	784
615	TTGGCCCTGCCCCCTTGCTCTGCT	GCCCTTAGATACAGCAGTTTATATACCAACACACCTGTGCT	674
785	ACACTGACATTCATAAAGTG	-ACGTGCTTGTGAAAAAAA	825
675	ACAGTGTCAATCAATAAGTGCAG	TGCTGTGTGAAAAAAA	716

RESULT 9	
CD176014	
LOCUS	726 bp mRNA linear EST 19-MAY-2003
DEFINITION	AGENCOURT_13962252 NIH_MGC_172 Homo sapiens cDNA 5', mRNA sequence.
ACCESSION	CD176014
VERSION	CD176014.1 GI:30861198
KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1 (bases 1 to 726)
AUTHORS	NIH-MGC <a href="http://mgc.nci.nih.gov/">http://mgc.nci.nih.gov/</a> .
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL	Unpublished (1999)
COMMENT	Contact: Robert Strausberg, Ph.D.

**FEATURES**  
**SOURCE**

RESULT 10  
CD175245  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1				
2				
3				
4				
5				
6				
7				
8				
9				
10				
11				
12				
13				
14				
15				
16				
17				
18				
19				
20				
21				
22				
23				
24				
25				
26				
27				
28				
29				
30				
31				
32				
33				
34				
35				
36				
37				
38				
39				
40				
41				
42				
43				
44				
45				
46				
47				
48				
49				
50				
51				
52				
53				
54				
55				
56				
57				
58				
59				
60				
61				
62				
63				
64				
65				
66				
67				
68				
69				
70				
71				
72				
73				
74				
75				
76				
77				
78				
79				
80				
81				
82				
83				
84				
85				
86				
87				
88				
89				
90				
91				
92				
93				
94				
95				
96				
97				
98				
99				
100				

CD175245 721 bp mRNA linear EST 19-MAY-2003  
 AGENCOURT\_13972682 NIH\_MGC\_172 Homo sapiens cDNA 5', mRNA sequence.  
 CD175245  
 CD175245.1 GI:30859628  
 EST.  
 Homo sapiens (human)  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 721)  
 NIH-MGC <http://mgc.nci.nih.gov/>.  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: [cgabbs-rc@mail.nih.gov](mailto:cgabbs-rc@mail.nih.gov)  
 Tissue Procurement: Dr. Jamie Thompson, University of WI

CDNA Library Preparation: Gina Zastrow-Hayes  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: NDKM41 row: C column: 22  
 High quality sequence start: 4  
 High quality sequence stop: 721.  
 Location/Qualifiers  
 1. 721  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /tissue\_type="embryonic stem cells, WA01, passage 38"  
 /lab\_host="DH10B TONa"  
 /clone\_lib="NIH\_MGC\_172"  
 /notes="Vector: pDONR201; Site 1: attP2; Site 2: attP1;  
 LIBR PRIMING - oligo dT; METHOD - full-length enriched;  
 Embryonic Stem Cells H1; LIBR PROVIDER - Bradfield"

ORIGIN  
 Query Match 60.6%; Score 671.8; DB 14; Length 721;  
 Best Local Similarity 98.8%; Pred. No. 7.7e-126;  
 Matches 687; Conservative 0; Mismatches 7; Indels 1; Gaps 1;  
 QY 141 CTGAGAACTGAAACGACAGCGGGAAGAGGTCTCACTGAGCAGCGTCCAGCATCCCGA 200  
 DB 20 CTGAGAACTGAAACGACAGCGGGAAGAGGTCTCACTGAGCAGCGTCCAGCATCCCGA 79  
 QY 201 CACCAAGCGGCGCTTGGTCCAGCGAGAAACACACACTTCTCAAACTTCACCTCAACAC 260  
 DB 80 CACCAAGCGGCGCTTGGTCCAGCGAGAAACACACACTTCTCAAACTTCACCTCAACAC 139  
 QY 261 TTCTTTCCCAAGCAGAGAGTGCACAGGAGGAATGAGTGGTGTCTGGGGCA 320  
 DB 140 TTCTTTCCCAAGCAGAGAGTGCACAGGAGGAATGAGTGGTGTCTGGGGCA 199  
 QY 321 CCCCCAGCACCATCTTCCAAAGGTCCACCGTGATCAACATCCACAGGAGACTTCCGTG 380  
 DB 200 CCCCCAGCACCATCTTCCAAAGGTCCACCGTGATCAACATCCACAGGAGACTTCCGTG 259  
 QY 381 CCGGACCATGTGTCGTGTCCTGTTCAACACCTCTTCTGACCTGTCGTGTCGGC 440  
 DB 260 CCGGACCATGTGTCGTGTCCTGTTCAACACCTCTTCTGACCTGTCGTGTCGGC 319  
 QY 441 TTCTAGCATTCGCTTACTCCGTGAGTCTAGGACAGGAAGTGGTGGGAGTGACC 500  
 DB 320 TTCTAGCATTCGCTTACTCCGTGAGTCTAGGACAGGAAGTGGTGGGAGTGACC 379  
 QY 501 GGGGCGCAGGCTATGCTTCCAGCGCAAGTGCCTGAACATCTGGGCGCTGATTCGGGC 560  
 DB 380 GGGGCGCAGGCTATGCTTCCAGCGCAAGTGCCTGAACATCTGGGCGCTGATTCGGGC 439  
 QY 561 ATCTCTATGACCATTCGATTCATCTCTTACTGTGATTCGGCTCTGTGACAGTCTACCAT 620  
 DB 440 ATCTCTATGACCATTCGATTCATCTCTTACTGTGATTCGGCTCTGTGACAGTCTACCAT 499  
 QY 621 ATTATGTTACAGATAATACAGGAAAAACGGGTTACTAGTAGCGGCCCATAGCCTGCAAC 680  
 DB 500 ATTATGTTACAGATAATACAGGAAAAACGGGTTACTAGTAGCGGCCCATAGCCTGCAAC 559  
 QY 681 CTTTGCACCTCCAGTGTGCAATGTGGCGCTGACAGCTGGGCTGTGGCGCTGCGCCCTT 740  
 DB 560 CTTTGCACCTCCAGTGTGCAATGTGGCGCTGACAGCTGGGCTGTGGCGCTGCGCCCTT 619  
 QY 741 GGTCTCTGCCCTAGATACAGAGTTTATACCCACACACCTGTCTACACTGTGACATTCATA 800  
 DB 620 GGTCTCTGCCCTAGATACAGAGTTTATACCCACACACCTGTCTACAGTGTGATTCATA 679  
 QY 801 AAGTG-ACGTGCTTGTGAAAAAATAAATAA 834  
 DB 680 AAGTGACGTGCTTGTGATAAAAAAATAAATAA 714

RESULT 11  
 AV759074

LOCUS 727 bp mRNA linear EST 19-OCT-2000  
 AV759074 MDS Homo sapiens cDNA clone MDSEPH09 5', mRNA sequence.  
 AV759074 DEFINITION  
 AV759074 ACCESSION  
 AV759074 VERSION  
 AV759074.1 GI:10916922  
 EST.  
 KEYWORDS  
 SOURCE  
 Homo sapiens (human)  
 ORGANISM  
 Homo sapiens

REFERENCE  
 AUTHORS

1 (bases 1 to 727)  
 Gu, J., Zhao, M., Huang, Q., Xu, X., Li, Y., Peng, Y., Song, H., Xiao, H.,  
 Gu, Y., Li, N., Qian, B., Liu, F., Qu, J., Gao, X., Cheng, Z., Xu, Z.,  
 Zeng, L., Xu, S., Gu, W., Tu, Y., Jia, J., Fu, G., Ren, S., Zhong, M.,  
 Lu, G., Yang, Y., Gao, G., Zhang, Q., Chen, S., Han, Z. and Chen, Z.  
 Homo sapiens cDNA MDS clones  
 Unpublished (2000)  
 Contact: Zeguang Han  
 Chinese National Human Genome Center at Shanghai  
 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai  
 201203, P. R. China  
 Tel: 86-21-50801919 (ex.45)  
 Fax: 86-21-50801922  
 Email: hanzg@chgc.sh.cn  
 This clone is available at CHGC in Shanghai.

TITLE  
 JOURNAL  
 COMMENT

FEATURES  
 source

1. 727  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="MDSEPH09"  
 /tissue\_type="Bone marrow"  
 /cell\_type="CD34+ hematopoietic stem/progenitor cell"  
 /lab\_host="BM25.8"  
 /clone\_lib="MDS"  
 /notes="Vector: pTriplex2; Site\_1: sfiIA; Site\_2: sfiIB"

ORIGIN

Query Match 60.3%; Score 668.8; DB 9; Length 727;  
 Best Local Similarity 98.4%; Pred. No. 3.1e-125;  
 Matches 695; Conservative 0; Mismatches 9; Indels 2; Gaps 2;  
 QY 129 AAATACACACTTCTGAGAA-ACGTGAAACGACAGGGGAAAGAGGTCTCACTGAGCACCCT 187  
 DB 1 AAATACACACTTCTGAGAGACTGAAACGACAGGGGAAAGAGGTCTCACTGAGCACCCT 60  
 QY 188 CCGAGATCCGGACACACAGCGGCCCTTCGCTCCAGCAGAGAAACCACTTCTCAAC 247  
 DB 61 CCGAGATCCGGACACACAGCGGCCCTTCGCTCCAGCAGAGAAACCACTTCTCAAC 120  
 QY 248 CTTCACTCAACACTTCTCTTCCCAAGCCAGACCATCTTCAAGGTCCACCGTGTCAACATCCACAG 367  
 DB 121 CTTCACTCAACACTTCTCTTCCCAAGCCAGAGATGCAAGAGGAGACATGAGGTGGC 180  
 QY 308 TGTGCTGGGGGACACCCCGCAGACCATCTCTTCAAGGTCCACCGTGTCAACATCCACAG 367  
 DB 181 TGTGCTGGGGGACACCCCGCAGACCATCTCTTCAAGGTCCACCGTGTCAACATCCACAG 240  
 QY 368 CGAGACCTCCGTGCGCGCCCATGTCGTGCTGCTTCAACACCCCTTCTTCTGAACCTG 427  
 DB 241 CGAGACCTCCGTGCGCGCCCATGTCGTGCTGCTTCAACACCCCTTCTTCTGAACCTG 300  
 QY 428 GTGCTGTCTGGGCTTCATAGCATTCGCTACTCCGTAAGTCTTAGGAGCAGGAAGATGGT 487  
 DB 301 GTGCTGTCTGGGCTTCATAGCATTCGCTACTCCGTAAGTCTTAGGAGCAGGAAGATGGT 360  
 QY 488 TGGCGACGTGACCGGGGCCCGCCCTATGCTCCACCGCCAAAGTGCCTGAACATCTGGGC 547  
 DB 361 TGGCGACGTGACCGGGGCCCGCCCTATGCTCCACCGCCAAAGTGCCTGAACATCTGGGC 420

```
QY 548 CCTGATCTGGGATCCTCATGACCAATGGATTATCTCTGTTACTGCTGTTATTCGGCTCTGT 607
Db 421 CCTGATCTGGGATCCTCATGACCAATGGATTATCTCTGTTACTGCTGTTATTCGGCTCTGT 480
QY 608 GACAGCTTACCATATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCGGCC 667
Db 481 GACAGCTTACCATATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCGGCC 540
QY 668 CATAGCCTGCAACCTTTGCACTCCACTGTGCAATGCTGCAAGCTGCGCCCTGCAAGCTGGGGCTGTTG 727
Db 541 CATAGCCTGCAACCTTTGCACTCCACTGTGCAATGCTGCGCCCTGCAAGCTGGGGCTGTTG 600
QY 728 CCCTGCCCCCTTCTGCTCTGCTGCTCTAGATACAGCAGTTTATACCCACACACCTGTCTACA 787
Db 601 CCCTGCCCCCTTCTGCTCTGCTGCTCTAGATACAGCAGTTTATACCCACACACCTGTCTACA 660
QY 788 CTGACATTCATTAAGTG-ACGTGCTTGTGAAAAAACAATAA 832
Db 661 GTGTCAATCAATAAGTGACGTGCTTGTGAAAAAACAATAA 706

RESULT 12
CD173064
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

CD173064 704 bp mRNA linear EST 19-MAY-2003
AGENCOURT_13974182 NIH_MGC_172 Homo sapiens cDNA 5', mRNA sequence.
CD173064.1 GI:30855257
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Tissue Procurement: Dr. Jamie Thompson, University of WI
cDNA Library Preparation: Gina Zastrow-Hayes
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: NDKM46 row: h column: 04
High quality sequence start: 4
High quality sequence stop: 704.
High quality sequence stop: 704.
Location/Qualifiers
1. 704
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="embryonic stem cells, WA01, passage 38"
/lab_host="DH10B Tona"
/clone_lib="NIH_MGC_172"
/notes="vector: pDONR201; Site1: attP2; Site2: attP1;
LIBR_PRIMING - oligo dt; METHOD - full-length enriched;
Embryonic Stem Cells H1; LIBR PROVIDER - Bradfield"

Query Match 60.2%; Score 668; DB 14; Length 704;
Best Local Similarity 98.4%; Pred. No. 4.5e-125;
Matches 68; Conservative 0; Mismatches 10; Indels 1; Gaps 1;

QY 143 GAGAACTGAACACAGGGAAGAGGTCTCACTGAGCACCCTCCAGCATCCGGACA 202
Db 9 GAGTAGAAGAACACAGGGAAGAGGTCTCACTGAGCACCCTCCAGCATCCGGACA 68
QY 203 CCACAGGGCCCTTGGCTCCAGCGGAGAAAAACACACTTCTCAACCTTCACTCAACACTT 262
Db 69 CCACAGGGCCCTTGGCTCCAGCGGAGAAAAACACACTTCTCAACCTTCACTCAACACTT 128
```

```
QY 263 CTTTCCCCAAGACGAGAGATGCAAGAGGAAACATGAGGTGGTGTGCTGGGGGCAAC 322
Db 129 CTTTCCCCAAGACGAGAGATGCAAGAGGAAACATGAGGTGGTGTGCTGGGGGCAAC 188
QY 323 CCCAGACACCATCTTCCAAAGTCCACCGTGATCAACATCCACAGCGAGACCTCCGTCGCC 382
Db 189 CCCAGACACCATCTTCCAAAGTCCACCGTGATCAACATCCACAGCGAGACCTCCGTCGCC 248
QY 383 CGACCATGTGCTCTGGTCCCTGTTCAACACCCCTCTTCTTTGAACTGTGTCTGGGCTT 442
Db 249 CGACCATGTGCTCTGGTCCCTGTTCAACACCCCTCTTCTTTGAACTGTGTCTGGGCTT 308
QY 443 CATAGCATTGSCCTACTCCGTGAGTCTAGGACAGGAAGATGGTTGGGACGCTGACCGG 502
Db 309 CATAGCATTGSCCTACTCCGTGAGTCTAGGACAGGAAGATGGTTGGGACGCTGACCGG 368
QY 503 GGCCAGGCGCTATGCTCCACCGCAAGTGCTGAACATCTGGGCGCTGATTTGGGCGAT 562
Db 369 GGCCAGGCGCTATGCTCCACCGCAAGTGCTGAACATCTGGGCGCTGATTTGGGCGAT 428
QY 563 CCTCATGACCATTTGGATTTCCTGTTACTGCTGTTATTCGGCTCTGTGACAGTCTACCATAT 622
Db 429 CCTCATGACCATTTGGATTTCCTGTTACTGCTGTTATTCGGCTCTGTGACAGTCTACCATAT 488
QY 623 TATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCGGCCCATAGCCTGCAACCT 682
Db 489 TATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCGGCCCATAGCCTGCAACCT 548
QY 683 TTGCATCCCACTGTGCAATGCTGGCCCTGCACGCTGGGGCTGTGSCCTGCCCCCTTGG 742
Db 549 TTGCATCCCACTGTGCAATGCTGGCCCTGCACGCTGGGGCTGTGSCCTGCCCCCTTGG 608
QY 743 TCCTGCCCCCTAGATACAGCAGTTTATACCCACACACCTGTCTACACTGACATTCATATAA 802
Db 609 TCCTGCCCCCTAGATACAGCAGTTTATACCCACACACCTGTCTACACTGACATTCATATAA 668
QY 803 GTG-ACGTGCTGTGAAAAAACAATAAATAAACC 837
Db 669 GTGCACGTGCTGTGAAAAAACAATAAATAAACC 704

CD175397 715 bp mRNA linear EST 19-MAY-2003
AGENCOURT_13971659 NIH_MGC_172 Homo sapiens cDNA 5', mRNA sequence.
CD175397
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

CD175397.1 GI:30859931
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Tissue Procurement: Dr. Jamie Thompson, University of WI
cDNA Library Preparation: Gina Zastrow-Hayes
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: NDKM41 row: n column: 17
High quality sequence start: 2
High quality sequence stop: 715.
High quality sequence stop: 715.
Location/Qualifiers
1. 715
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"

FEATURES
source
```



